BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Plant Molecular Biology 24: 853–862, 1994.

A cereal haemoglobin gene is expressed in seed and root tissues under anaerobic conditions

Euan R. Taylor[1], Xian Zhou Nic[1], Alexander W. MacGregor[2] and Robert D. Hill[1,*]
[1] *Department of Plant Science, University of Manitoba, Winnipeg, Manitoba, Canada, R3T 2N2 (*author for correspondence);* [2] *Grain Research Laboratory, Canadian Grain Commission, Winnipeg, Manitoba, Canada, R3C 3G8*

Received 25 October 1993; accepted in revised form 10 December 1993

Key words: aleurone, anaerobiosis, barley, haemoglobin, RNA

Abstract

Legumes, and a very few non-legume plant species, are known to possess functioning haemoglobin genes. We describe here the characterization of a haemoglobin cDNA isolated from barley. The deduced amino acid sequence shows 71% amino acid identity with a non-legume haemoglobin gene, a further 16% of the residues being conservative replacements. The barley cDNA also hybridizes to genomic sequences in rye, maize and wheat. The demonstration of a gene from a monocotyledon with close sequence homology to the known non-legume plant haemoglobins fills a major gap in the known distribution of haemoglobin genes in the plant kingdom. The expression of the gene is induced in isolated barley aleurone layers exposed to anaerobic conditions, and the roots of flooding-stressed barley plants. The expression of the RNA under anoxic conditions is similar to that of a known anaerobic response gene, alcohol dehydrogenase. Our results suggest that the increased expression of haemoglobin RNA is an integral part of the normal anaerobic response in barley. The findings are discussed in the light of current theories of haemoglobin function and evolution.

Introduction

Plant haemoglobins (Hb) in the form of leghaemoglobins (Lb) have, for many years, been known to exist in the root nodules of legumes [27, 30]. Haemoglobins have more recently been positively identified in two closely related members of the Ulmaceae *Parasponia andersonii* (nodulating) and *Trema tomentosa* (non-nodulating), and in one species *Casuarina glauca* of the Casuarinaceae [8, 29]. There is considerable support [8, 10, 14, 31] for the hypothesis that plant and animal Hbs have a common evolutionary origin. However, the inability to demonstrate the presence of an Hb gene in monocotyledons has left open the possibility of inter-kingdom horizontal transfer to the dicots [8, 31]. There have been no reports which unequivocally demonstrate the presence of an Hb gene in a monocotyledon [3, 6], although claims of Lb-like sequences in the genomes of a number of non-legumes have appeared [36, 37]. These, however, were based only on Southern blotting

The nucleotide sequence data reported will appear in the GenBank Nucleotide Sequence Database under the accession number U01228.

experiments and the results have been questioned [3]. Several authors [8, 31] have noted the limitations which have frustrated attempts at cross species screening using available antibody and DNA probes.

The role of Lb in legume symbioses has been well-documented [2]. It is believed to act as an oxygen carrier in the symbiosomes of root nodules, supplying oxygen to the bacterial respiratory chain whilst preserving a low free-oxygen concentration [2]. The sequences of other plant Hb proteins are quite distinct from those of the Lb type and have been reviewed in some detail by Arredondo-Peter and Escamilla [6]. Current knowledge of the properties of non-legume plant Hb is based largely upon studies of the *Parasponia* protein. This protein has a very high oxygen affinity (Km 89 nM), and shows rapid oxygen rebinding (geminate reaction) [17, 43]. Transgenic studies have shown that the *Parasponia* Hb promoter is activated in the root meristem and vascular bundle in tobacco, and additionally expressed in root nodules of *Lotus* [9, 31].

The main body of work on the response of plants to anaerobiosis has dealt with glycolytic metabolism. Seed and root tissues that are tolerant of hypoxic conditions respond to oxygen deficit by increased utilization of glycolytic pathways [28]. In barley aleurone tissue, for example, low oxygen stress results in increased expression of alcohol dehydrogenase and lactate dehydrogenase RNA and protein [20, 22, 26]. Whilst changes in glycolysis are clearly major factors in plant adaptation to hypoxia, there are probably other factors involved. It is likely that under otherwise anoxic conditions roots can receive some oxygen input by axial movement from the aerial parts of the plant [4, 5, 11]. It has also been suggested that facilitated diffusion of oxygen and the accompanying buffering of oxygen tensions may be significant factors for cellular metabolism [40].

We have isolated a barley cDNA clone which has strong sequence homology to other non-legume Hbs. We have examined the Hb transcript levels, under varying exposures to oxygen stress, in comparison to the response of alcohol and lactate dehyrodegenase, two glycolytic enzymes that have been studied extensively in the normal anaerobic response of aleurone tissues. We also examined the behaviour of the gene in root tissues under flooding (hypoxic) conditions. On the basis of our results, we suggest that all monocots, and probably all plants, may in fact possess an Hb gene. Our evidence also indicates that the expression of the barley Hb gene is highly induced in response to anaerobic stress, and that this property of the gene is not limited to barley. The implications of these findings are discussed.

Materials and methods

Materials

Wheat (cv. Biggar), rye (cv. Musketeer) and maize were available locally. Barley seeds (cv. Harrington) were provided by the Canadian Grain Commission, Winnipeg, Manitoba. The Adh 1 probe was a gift from E. S. Dennis, and the barley Ldh probe was a gift from A. D. Hanson.

Library screening and sequence analysis

Clones were isolated from a λgt22 cDNA library of RNA from aleurone layers of Harrington barley half seeds, incubated with 1 μM gibberellic acid for 5 days [11, 34]. The initial Hb cDNA was isolated serendipitously in an immunological screen for other proteins. This cDNA was then used as a probe to identify full-length clones. The library was screened with the radiolabelled Hb cDNA probe according to Maniatis *et al.* [32]. cDNA inserts were subcloned into bluescript SK plasmid, and sequenced using a Gibco/BRL cycle sequencing kit. Sequence analysis and manipulation were carried out using the FASTA and FSAP packages and other software [15, 16, 35].

Southern blotting

Genomic DNA samples were prepared from seedling leaves by a modification of the CTAB

method [38] and aliquots of 10 or 20 μg DNA were digested with appropriate restriction enzymes before size fractionation on agarose gels in Tris-acetate buffer. DNA was blotted onto Hybond N^+ (Amersham) membrane in 20 × SSC overnight and fixed by alkali and UV treatment. Hybridisations were carried out overnight in a mixture containing 6 × SSC, 5 × Denhardt's solution, 1% SDS, 100 μg/μl herring testis DNA, at 65 °C. Washes were carried out at 65 °C in 0.5% SDS plus varying concentrations of SSC. Washes were initially performed in 2 × SSC, then in 1 × SSC, 0.5 × SSC, or finally 0.2 × SSC as appropriate. Filters were exposed to Kodak XAR-5 film with screens at −75 °C. Fragment sizes were estimated relative to Gibco/BRL size standards.

Northern blotting

Total RNA was prepared from frozen seedling leaves, seedling roots, coleoptiles of 10 day dark germinated seeds, and aleurone layers (usually batches of 75 aleurone layers per sample) according to the procedure described by Mohapatra *et al.* [33]. RNA was quantitated by UV absorption spectroscopy [32] and size-fractionated on agarose/formaldehyde gels [32]. Loading equivalence and RNA integrity were confirmed by ethidium staining of gels prior to transfer. RNA was blotted to GeneScreen (NEN) or Hybond N^+ (Amersham) membranes according to the manufacturer's instructions, and fixed with UV. Hybridisation, washing, and exposure to film, were carried out as described for southern blotting. Transcript sizes were estimated relative to Gibco/BRL RNA size standards, which were visualised by methylene blue staining [32]. Blots were re-probed in sequence with Hb probe, Ldh probe and Adh probe.

Preparation and use of probes

Purified plasmid inserts were radioactively labelled according to standard procedures [32]. Labelling was performed using ^{32}P-dCTP (ICN), unincorporated label was removed by spinning through Sephadex G50 (Pharmacia). Probes were added at concentrations of $1–5 \times 10^6$ cpm per ml of hybridisation fluid. Higher probe concentrations were used for cross-species Southern blots where signal strengths were extremely low.

Preparation of aleurone layers

Aleurone layers from barley (*Hordeum vulgare* cv. Harrington) were isolated by normal procedures [11, 34]. Briefly, embryo-free barley half seeds were sterilised in 0.25% bleach for 30 min, rinsed four times for 5 to 10 min in sterile water. After 3 days imbibition at 22 °C, aleurone layers were separated. Twenty or 25 layers were placed in sterile 50 ml conical flask containing 3 ml H_2O, 1 μM GA, 10 μg/ml chloramphenicol, 10 mM $CaCl_2$, and incubated on a rotary shaker running at 65 cycles/minute, at ca. 24 °C.

Oxygen stress treatments

Isolated aleurone layers were incubated under controlled oxygen tensions essentially as described by Hanson and Jacobsen [22]. For time course incubation under N_2, each flask was flushed with high purity N_2 for 2 min, then flasks were placed in 3 litre jars, which were in turn purged with N_2 for 1 h. The jars were reflushed with N_2 for 30 min every 6 h thereafter. For varied oxygen tensions, aleurone flasks were placed in 1-litre jars; these were sealed and flushed with the appropriate mixture of O_2 and N_2 gases for 15 min. Jars were re-purged with the same gas mixture every 3 to 6 h. Samples were harvested at appropriate time points, frozen in liquid nitrogen, and stored at −75 °C.

Flooding experiments

Batches of 10 barley or maize seeds were planted in a mixture of peat, sand, and soil and allowed

to germinate normally up to the two leaf stage. Individual pots were used for each treatment. At the start of the experimental treatment, an appropriate number of pots were placed in a deep tray of water, so as to completely submerge the root system up to soil level. For sampling, the contents of whole pots were removed and washed to clear the roots of soil and debris. Root tissue was frozen in liquid nitrogen and used to prepare total RNA as described above.

Extraction and assay of ADH and LDH

Batches of 6 aleurone layers were homogenised using a blender (Polytron, Brinkman) at top speed for 30 s in a 10 ml plastic tube containing 2 ml ice-cold isolation buffer (0.5 M Tris-HCl pH 8.0, 10 mM DTT). The homogenate was clarified by centrifugation at 10000 rpm (12000 × *g*) in a Sorvall SS34 rotor for 20 min. The supernatants were transferred to 1.5 ml microfuge tubes and respun at 10000 rpm for 20 min. The supernatants were used for assays. Lactate dehydrogenase (EC 1.1.1.1) and alcohol dehydrogenase (EC 1.1.1.27) assays were performed exactly as described by Hanson and Jacobsen [22]. Total protein concentrations were standardised using a Biorad protein assay kit with γ-globulin as a standard. Enzyme activity was expressed in IU (μmol of NADH oxidised or reduced per minute). Each time point represents the results of three separate measurements.

Results

Cloning and sequencing

During the course of immunologically screening an aleurone cDNA library with antiserum for the 1,6-α-glucosidase, limit dextrinase, a cDNA was isolated which encoded a protein with strong sequence homology to *Parasponia andersonii* and *Trema tomentosa* Hbs [8, 31]. It is possible that this is due to some contaminant in the protein used to raise antiserum, but the serum does not detect an Hb-sized protein on western blots. The initial cDNA isolate represents bases 354 to 706 of the sequence presented here (Fig. 1). This fragment was used as a probe to re-screen the library yielding a number of clones of varying lengths, the sequence of the longest clone is the basis for the data presented. Partial sequencing of a number of independent clones indicated no rearrangement of the cDNA sequence during cloning. It also revealed no variations in nucleotide sequence between different clones.

The complete open reading frame encodes a predicted translation product 162 amino acids long (17.8 kDa). The predicted protein sequence shows 71% amino acid identity with *Parasponia* Hb, with a further 16% of residues being conservative replacements. The sequence is clearly distinct from the Lb class of proteins, relative to which it shows ca. 40% sequence identity and a further 40% of conserved changes. Residues conserved between known dicotyledons and animal Hb [6] are also conserved in the barley Hb. These residues are underlined in Fig. 1. The nucleotide sequence is likewise highly conserved relative to other non-legume Hbs, being 70% homologous to *Parasponia*. Sequence comparisons of different plant Hbs have revealed a consensus for nine specific residues amongst those which show rapid oxygen rebinding (geminate plus) [17], six of these are also conserved in the barley and *Parasponia* sequences. Two of the changed positions (Thr-92 and Gln-149) are identical in the barley and *Parasponia* sequences (see Fig. 1). The barley protein also shares the cysteine residue (residue 79 in the barley protein) seen in all the sequenced non legume plant Hbs and in none of the sequenced Lbs [6]

The haemoglobin DNA sequence is present in the genomic DNA of barley and other cereals

Southern blotting (Fig. 2a) confirmed that the cDNA sequence was present in the barley genome. Hybridising fragments were ca. 5.0 kb with *Bam* HI, 4.9 kb with *Bgl* II, 7.5 kb with *Hind* III, and 4.0 kb with *Bam* HI and *Hind* III combined.

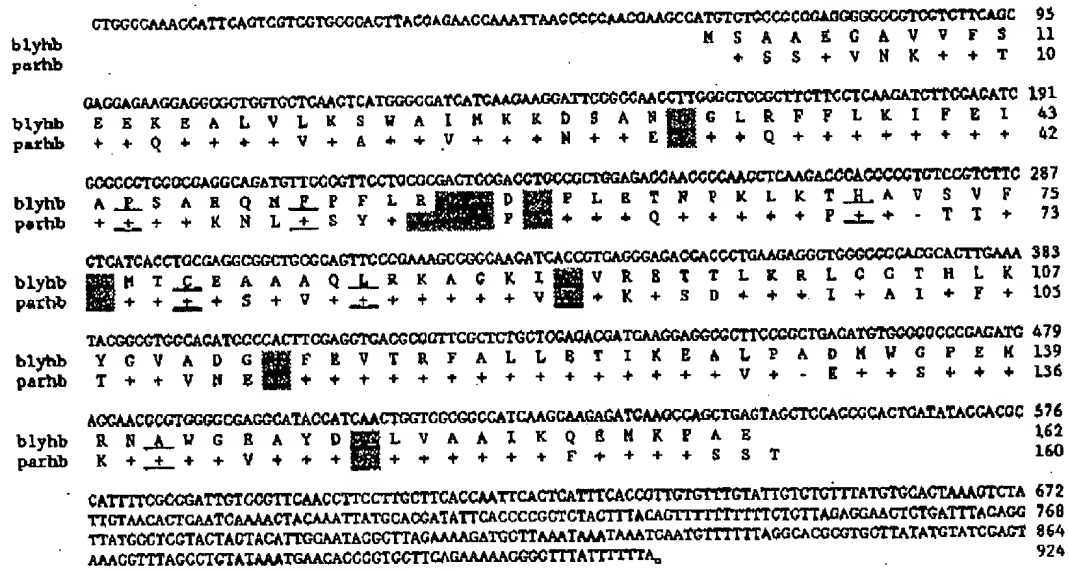

Fig. 1. Nucleotide sequence of a barley Hb cDNA clone with predicted translation product (blyhb), showing similarity to *Parasponia* Hb [8] (parhb). + indicates identical residue, - indicates gap inserted to optimise similarity. Amino acid residues conserved between animal and dicot Hbs [6] are underlined. Positions of the nine residues conserved amongst the geminate plus (fast oxygen rebinding) Hbs are shaded. Thr-92 and Gln-149 replace Leu and Val, respectively, in the geminate plus consensus (double-underlined) [17]. The conserved cysteine (Cys-79) is also indicated (double-underlined). Nucleotide and protein sequences are numbered.

The results suggest that in barley, like in *Parasponia* [8], the gene is present at very low copy number. Lower-stringency washes did not indicate the presence of multiple, related sequences. None of the blotting results suggest the presence of a multigene family as seen with the Lbs.

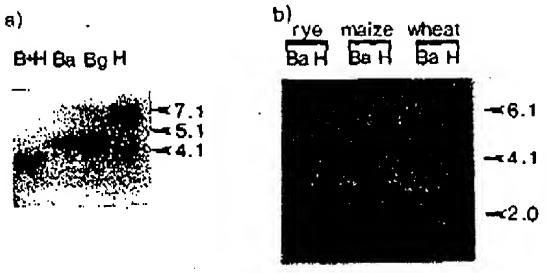

Fig. 2. Southern blots probed with barley Hb cDNA. a. Barley DNA digested with, *Bam* HI and *Hind* III (B + H), *Bam* HI (Ba), *Bgl* II (Bg), *Hind* III (H). b. Rye, maize and wheat DNA, digested with *Bam* HI (Ba) or *Hind* III (H). Samples were run on a 1% agarose gel and blotted as described in Materials and methods. Positions of size markers are indicated in kb. Unhybridised probe was removed by washing in 2× SSC and then 0.5 × SSC/0.5%; SDS (rye, maize, wheat) or 0.2 × SSC/0.5%; SDS (barley), at 65 °C. Blots were exposed to film at −75 °C.

The barley probe also hybridised to genomic sequences in a number of other cereals (Fig. 2b), even at quite high stringency. In rye, *Bam* HI gave a 4.8 kb hybridising band, and *Hind* III, 7.5 kb and 5.0 kb bands. In maize, *Bam* HI gave a fragment of 4.8 kb and *Hind* III 7.5 kb. Within the limits of cross-species probing, it seems that in both of these cases, as in barley and *Parasponia*, the gene is present at low copy number. Again, even at low stringency no evidence was seen for a multigene family as observed amongst the Lbs. In hexaploid wheat, there are three hybridising fragments with each enzyme, *Bam* HI giving bands at 6.3, 2.4 and 1.8 kb, and *Hind* III giving bands of 7.6, 5.0 and 3.8 kb. It is possible that each of the three bands represents one chromosome complement. We cannot rule out the possibility, however, that these three bands represent a gene family.

The haemoglobin RNA is expressed in roots and aleurone layers

Northern blot analysis of barley RNA indicated a transcript of ca. 900 nucleotides, which was present in both total and poly(A)$^+$ RNA (Fig. 3a and b). A transcript was easily detectable in total RNA isolated from the roots of seedlings at the one- and three-leaf stages. There was no detectable expression of the transcript in leaf tissue of the same plants. This is consistent with the expression pattern seen in the non-legume dicot species which have been examined [3, 8].

In addition to its expression in root tissue, the Hb transcript can be detected at very low levels in coleoptile tissue (Fig. 3a) and in oxygen-stressed aleurone layers (Fig. 3b) of imbibed half seeds.

Effect of O_2 concentration on the expression of haemoglobin, alcohol dehydrogenase and lactate dehydrogenase mRNA expression

Since the expression of Hb message increased in response to O_2 deprivation, a series of experiments were conducted in which aleurone layers were incubated under imposed oxygen concentrations ranging from 20% to 0% O_2 for a period of 24 h. Hb mRNA was not detectable at 20% and 10% O_2 was barely visible at 5% O_2 and became easily detectable at 2% and 0% O_2 (Fig. 4a). The same blots used to examine the Hb message were re-probed for Adh and Ldh transcripts. Adh transcript levels increased steadily as the O_2 concentration was decreased from 10% to 0%. Ldh message was noticeably elevated between 2% and 10% O_2 with a lower but easily detectable basal level of expression in 0% and 20% O_2.

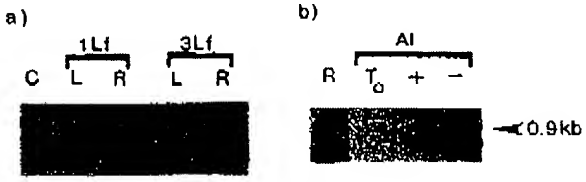

Fig. 3. Northern blots of barley RNA probed with barley Hb cDNA. a. 30 μg total RNA from, coleoptiles (C), leaves (L), and roots (R). Growth stages of 1 leaf and 3 leaf are indicated. b. 0.5 μg poly A$^+$ RNA from roots of 3-leaf stage seedlings (R). 20 μg total RNA from aleurone layers (Al) at a zero time point (T_0) and after 24 h in normal atmospheric conditions (+), or oxygen deficient conditions (−). Approximate transcript size is shown in kb. Gels were run and blotted as described in Materials and methods. Blots were washed in 0.2× SSC at 65 °C.

Temporal expression of haemoglobin, alcohol dehydrogenase and lactate dehydrogenase RNA in aleurone layers under anoxic conditions

The expression of Hb, Ldh and Adh messages was examined in aleurone layers under sustained anoxia (Fig. 4b). Increased levels of Hb message were first observable after 1 h of incubation under N_2, and continued to increase up to 12 h, the last sampling point in this experiment. No significant changes in Hb message level were observed beyond 12 h (results not shown). Increases in Adh and Ldh transcript levels also became apparent after the first hour of anoxia. After the first two hours of treatment changes became less pronounced. High basal levels of Ldh message were apparent in the 0 h and 12 h control (incubated in normal air) samples.

Changes in enzyme activity of ADH and LDH under anoxia and hypoxia

As a confirmation that the aleurone preparations were behaving normally under oxygen stress, activity assays for LDH and ADH enzymes were carried out on aleurone layers exposed to a range of imposed O_2 concentrations, and under sustained anoxia. When aleurones were incubated under O_2 concentrations ranging from 20% to 0% (Fig. 4a, top), the results indicated that ADH activity remained low until the O_2 concentration dropped below 10%, activity then increased steadily up to a maximum in 0% O_2 LDH activity rose as the O_2 concentration was decreased to

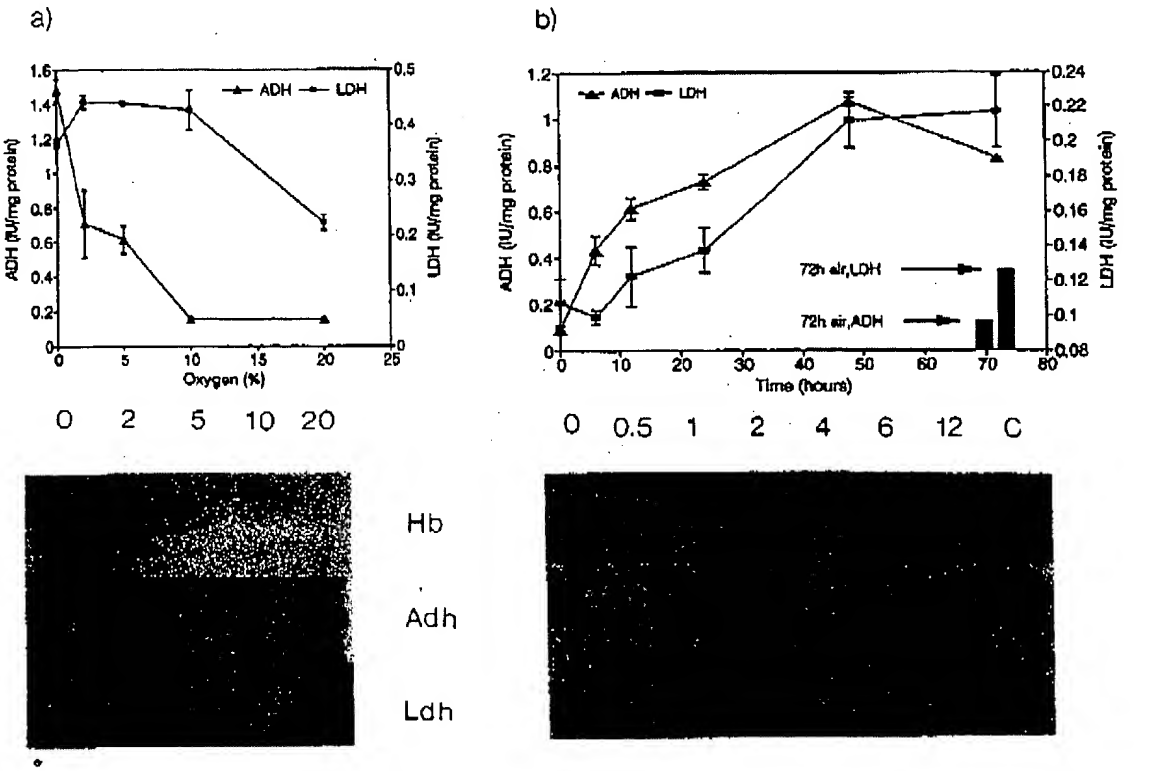

Fig. 4. Response of aleurone layers to (a) different oxygen concentrations, and (b) sustained anoxia. Northern blots were probed in sequence with Hb probe, Ldh probe, and Adh probe. Each lane carries 15 μg total RNA, blots were washed at 0.2 × SSC 65 °C. In a, numbers represent percentage oxygen, from 0 to 20%. In b, numbers represent time in hours after beginning the experiment, from 0 to 12 hours. Graphs show the corresponding changes in ADH and LDH enzyme activity observed in similar experiments. The standard deviation for each measurement is shown. The time-course of enzyme activities shown in b is over 72 h to allow direct comparison with the results of Hanson *et al.* [22]. Gels were run and blotted as described in Materials and methods. Blots were washed in 0.2 × SSC at 65 °C.

10%, remained relatively constant over the range of 10% to 2% O_2, then declined slightly at 0% O_2.

Under sustained anoxia (Fig. 4b, top) ADH activity increased rapidly during the first 12 h, and continued to increase up to 48 h. Following a slight decline in the first 6 h, LDH activity increased steadily until 48 h. A comparison of anoxic and oxygenated aleurone layers after 72 h indicated about an 8-fold increase in ADH activity and less than a 2-fold increase in LDH activity.

These results are very similar to those of Hanson and Jacobsen [22]. They confirm that the aleurone preparations behaved normally in response to low oxygen stress. They place results regarding Hb transcription in the context of a normal anaerobic response.

Levels of the haemoglobin transcript are increased in flooded roots of barley and maize

In view of the increase in Hb transcript levels in oxygen-stressed aleurone layers it was important to know if this finding was also relevant to the gene in its major site of expression, root tissue. Barley and maize roots were examined for Hb RNA expression under flooding stress. Figure 5 shows RNA samples from the flooded and unflooded plants, probed with the Hb cDNA. After a 12 h period of flooding stress, both maize and

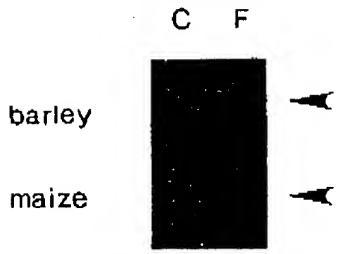

Fig. 5. Hb transcript expression in RNA from the roots of barley and maize plants. Lanes are (C) control roots, and (F) flooded root systems, completely submerged in water for 12 h. Each lane carries 15 μg of total RNA. Gel was run and blotted as described in Materials and methods. Blot was washed in 0.2 × SSC at 65 °C.

barley expressed greatly increased levels of Hb RNA. In this exposure the Hb transcript is barely visible in control roots, but is clearly visible in the flooding-stressed samples.

Discussion

It is now generally accepted that the known animal and dicot plant Hbs evolved from a common ancestor about 1400 million years ago [10, 14, 31]. However, in the absence of a monocot Hb gene, the possibility that the dicot gene is a result of interkingdom horizontal transfer has remained troubling [3]. The demonstration of a cereal Hb gene confirms that it is present in two of the major divisions of the plant kingdom. This, in turn, makes it all the more likely that an Hb gene is indeed present in all plants [3, 6, 14], and perhaps serves a common function there [3]. The conservation of kinetically important residues between the barley and *Parasponia* Hbs (Fig. 1) suggests that the barley protein is functional and that it may share kinetic properties with the *Parasponia* protein. Some authors have suggested that the identification of a monocot Hb gene may circumvent one conceptual obstacle in the generation of artificial nitrogen-fixing symbioses involving monocot plants [3].

Based on the available information concerning the expression pattern and kinetic properties of non-legume plant Hb, two major hypotheses have been put forward to account for the expression of the protein in non-nodulating plants. One involves Hb-facilitated diffusion of oxygen within the root tissues [3, 9]. Another proposes a role as an oxygen sensor, detecting changing levels of environmental oxygen and causing the switch from aerobic to anaerobic metabolism [3, 9]. There is already evidence from work on mammalian and rhizobial systems [18, 19] that at least some oxygen sensors are haem proteins. However, work on the response of *E. coli* to low oxygen stress suggests that the switch mechanism need not rely on monitoring oxygen levels at all [39]. Work on a range of experimental systems has suggested a number of other functions for proteins of the Hb family, including reduction of metal ions [7], and a role as a terminal oxidase [12, 13]. These and many other possibilities have been extensively reviewed by Wittenberg and Wittenberg [42].

Whether the major role of plant Hb is as an oxygen sensor, or as a facilitator of oxygen diffusion, the expression of the gene would be expected in organs such as roots and seed tissues (the aleurone and embryo itself) which might be expected to experience oxygen stress [3, 22]. Aside from one unconfirmed report of an Hb protein in seed tissues [25], there have been no previous reports of Hb expression other than in root and root nodule tissues.

If the Hb protein has some functional role in the anaerobic response (as a facilitator of oxygen diffusion for example), one would expect it's RNA levels to be up-regulated under low oxygen stress. That is, it should behave in much the same way as other genes involved in the anaerobic response [1, 22, 24, 41]. Experimental analysis showed that if barley aleurone layers are subjected to a low oxygen regime (Fig. 4), the level of the Hb RNA does indeed increase dramatically. The increase in transcript levels in response to low oxygen levels is apparent at lower O_2 concentrations for Hb than it is for Adh or Ldh. The timing of the Hb response to anoxia is similar to that of the Adh and Ldh genes, the increase in Hb transcript levels seems to be relatively large. Taken together with the changes in ADH and LDH enzyme activity, these results indicate that the induction of

Hb RNA expression is occurring within the framework of a normal anaerobic response [20, 22, 23]. The Hb RNA induction parallels rather than anticipates the expression of the Adh and Ldh messages. This would not be expected if the only function of the protein were as an oxygen sensor triggering the anaerobic response and thus the increased expression of these two transcripts.

An examination of the expression of the Hb RNA in roots of both barley and maize gives a similar picture, the gene is again very responsive to low oxygen stress. The results demonstrate that this responsiveness to O_2 levels is not limited to the aleurone system. It is also important to note that the Hb gene shows same O_2 responsiveness in at least two different monocot species.

The expression of an Hb gene in roots and aseptically cultured aleurone layers of a monocot with no known symbiotic partners strengthens the proposition that plant Hbs do have a more universal function than their role in symbioses [3, 31, 42]. If this is true, together with the high level of sequence conservation between the barley and dicot non-legume Hbs, it provides some circumstantial support for the proposal that even legumes may still possess a copy of the 'non-Lb' gene, from which the Lbs themselves have arisen for specialist purposes by gene duplication and subsequent evolution [3, 31]. Such a conserved locus could easily have been overlooked due to the well documented problems inherent in cross species screening with available Hb probes [3, 31]. In our own work, we have observed that Lb probes and barley Hb probes do not cross hybridise even at very low stringency and in the presence of a large excess of the target sequence and the probe (unpublished data).

Hb transcription in a well characterised cell system, the barley aleurone layer, provides a sufficiently defined system for transfection experiments to directly address the question of Hb function and its possible role in the anaerobic response [21, 22]. A clearer picture of this response is crucial to a better understanding of the factors which lead to flooding tolerance in some cereals and not in others [24]. We are currently pursuing experiments designed to directly assess the physiological significance of the Hb protein.

Acknowledgements

This work was funded by A Research Grant OGP4689 from the Natural Sciences and Engineering Research Council of Canada. We thank Doug Durnin for technical assistance and Ma Luo for the aleurone cDNA library. X.Z.N and E.R.T contributed equally to this work.

References

1. Andrews DL, Cobb BG, Johnson JR, Drew MC: The anaerobic response of maize seedlings differs greatly with seedling age. Plant Physiol 101: 407–414 (1993).
2. Appleby CA: Leghemoglobin and *Rhizobium* respiration. Annu Rev Plant Physiol 35: 443–478 (1984).
3. Appleby CA, Bogusz D, Dennis ES, Trinick MJ, Peacock WJ: A role for haemoglobin in all plant roots? Plant Cell Environ 11: 359–367 (1988).
4. Armstrong W: Aeration in Roots. In: Cherry JH (ed) Environmental Stress in Plants, pp. 197–206. NATO ASI Series vol G19. Springer-Verlag, Berlin/Heidelberg (1989).
5. Armstrong W, Beckett PM: Internal aeration and development of stelar anoxia in submerged roots: a multishelled mathematical model combining axial diffusion of oxygen in the cortex with radial losses to the stele, the wall layers and the rhizosphere. New Phytol 105: 221–245 (1987).
6. Arredondo-Peter R, Escamilla E: A concensus sequence of plant haemoglobins. Plant Mol Biol Rep 9: 195–207 (1991).
7. Bakan DA, Saltman P, Theriault Y, Wright PE. Kinetics and mechanisms of reduction of Cu(II) and Fe(III) complexes by soybean leghemoglobin α. Biochim Biophys Acta 1079: 182–196 (1991).
8. Bogusz D, Appleby CA, Landsmann J, Dennis ES, Trinick MJ, Peacock WJ: Functioning haemoglobin genes in non-nodulating plants. Nature 331: 178–180 (1988).
9. Bogusz D, Llewellyn DJ, Craig S, Dennis ES, Appleby CA, Peacock WJ: Nonlegume hemoglobin genes retain organ-specific expression on heterologous transgenic plants. Plant Cell 2: 633–641 (1990).
10. Brown GG, Lee JS, Brisson N, Verma DP: The evolution of a plant globin gene family. J Mol Evol 21: 19–32 (1984).
11. Chrispeels MJ, Varner JE: Gibberellic acid-enhanced synthesis and release of α-amylase and ribonuclease by

isolated barley aleurone layers. Plant Physiol 42: 398–406 (1967).
12. Dikshit RP, Dikshit KL, Liu Y, Webster DA: The bacterial hemoglobin from *Vitreoscilla* can support the aerobic growth of *Escherischia coli* lacking terminal oxidases. Arch Biochem Biophys 293: 241–245 (1992).
13. Dikshit KL, Webster DA: Cloning, characterisation and expression of the bacterial globin gene from *Vitreoscilla* in *Escherischia coli*. Gene 70: 377–386 (1988).
14. Dixon B, Pohajdak B: Did the ancestral globin gene of plants and animals contain only two introns? Trends Biochem Sci 17: 486–488 (1992).
15. Fristensky B: Improving the efficiency of dot-matrix similarity searches through the use of an oligomer table. Nucl Acids Res 14: 597–610 (1986).
16. Fristensky B, Lis J, Wu R: Portable microcomputer software for nucleotide sequence analysis. Nucl Acids Res. 10: 6451–6463 (1982).
17. Gibson QH, Wittenberg JB, Wittenberg BA, Bogusz D, Appleby C: The kinetics of ligand binding to plant hemoglobins. J Biol Chem 264: 100–107 (1989).
18. Gilles-Gonzalez MA, Ditta GS, Helinski DR: A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*. Nature 350: 170–172 (1991).
19. Goldberg M, Dunning SP, Bunn HF: Regulation of the erythropoietin gene: Evidence that the oxygen sensor is a heme protein. Science 242: 1412–1415 (1988).
20. Good AG, Crosby WL: Induction of alcohol dehydrogenase and lactate dehydrogenase in hypoxically induced barley. Plant Physiol 90: 860–866 (1989).
21. Gopalakrishnan B, Sonthayanon B, Rahmatullah R, Muthukrishnan R: Barley aleurone layers as a transient expression system. Plant Mol Biol 16: 463–467 (1991).
22. Hanson AD, Jacobsen JV: Control of lactate dehydrogenase, lactate glycolysis, and α-amylase by O_2 deficit in barley aleurone layers. Plant Physiol 75: 566–572 (1984).
23. Hanson AD, Jacobsen JV, Zwar JA: Induction of alcohol dehydrogenase and lactate dehydrogenase in barley aleurone layers. Plant Physiol 75: 573–581 (1984).
24. Harberd NP, Edwards KJR: The effect of a mutation causing alcohol deficiency on flooding tolerance in barley. New Phytol 90: 631–644 (1982).
25. Hirano H: Microsequence analysis of winged bean seed proteins electroblotted from two-dimensional gel. J Prot Chem 8: 115–130 (1989).
26. Hoffman NE, Bent AF, Hanson AD: Induction of lactate dehydrogenase isozymes by oxygen deficit in barley root tissue. Plant Physiol 82: 658–663 (1986).
27. Keilin FRS, Wang YL: Haemoglobin in the root nodules of leguminous plants. Nature 155: 227 (1945).
28. Kennedy RA, Rumpho ME, Fox TC: Anaerobic metabolism in plants. Plant Physiol 100: 1–6 (1992).
29. Kortt AA, Inglis AS, Fleming AJ, Appleby CA: Amino-acid sequence of hemoglobin-I from root-nodules of the non-leguminous *Casuarina glauca-Frankia* symbiosis. FEBS Lett 231: 341–346 (1988).
30. Kubo H: Hemoprotein from the root nodules of legumes. Acta Phytochim 11: 195 (1939).
31. Landsmann J, Dennis E, Higgins TJV, Appleby CA, Kortt A, Peacock WJ: Common evolutionary origin of legume and non-legume plant haemoglobins. Nature 324: 166–168 (1986).
32. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989).
33. Mohapatra SS, Poole RJ, Dhindsa RS: Changes in protein patterns and translatable messenger RNA populations during cold acclimation of alfalfa. Plant Physiol 84: 1172–1176 (1987).
34. Murthy PPN: A fast and easy technique for the isolation of aleurone layers. Plant Physiol 90: 388–389 (1989).
35. Pearson WP, Lipman DJ: Improved tools for biological sequence analysis. Proc Natl Acad Sci USA. 85: 2444–2448 (1988).
36. Roberts MP, Jafar S, Mullin BC: Leghaemoglobin-like sequences in the DNA of non-leguminous plants. In: Galau GA (ed) Abstracts First International Conference on Plant Molecular Biology p. 14. (University of Georgia, Athens 1985).
37. Roberts MP, Jafar S, Mullin BC: Leghaemoglobin like sequences in the DNA of four actinorhizal plants. Plant Mol Biol 5: 333–337 (1985).
38. Rogers SO, Bendich AJ: Extraction of DNA from plant tissues. In: Gelvin SA, Schilperoort RA (eds), Plant Molecular Biology Manual, pp. A6/1–A6/10. Kluwer Academic Publishers, Dordrecht (1991).
39. Spiro S, Guest JR: Adaptive responses to oxygen limitation in *Escherichia coli*. Treds Biochem Sci 16: 310–314 (1991).
40. Stokes AN: Facilitated diffusion: the elasticity of oxygen supply. J Theor Biol 52: 285–97 (1975).
41. Thomson CJ, Greenway H: Metabolic evidence for stelar anoxia in maize roots exposed to low O_2 concentrations: Plant Physiol 96: 1294–1301 (1991).
42. Wittenberg JB, Wittenberg BA: Mechanisms of cytoplasmic haemoglobin and myoglobin function. Annu Rev Biophys Biophys Chem 19: 217–241 (1990).
43. Wittenberg JB, Wittenberg BA, Gibson QH, Trinick MJ, Appleby CA: The kinetics of the reactions of *Parasponia andersonii* hemoglobin with oxygen, carbon monoxide, and nitric oxide. J Biol Chem 261: 13624–13631 (1986).

Plant Physiol. (1997) 115: 1259–1266

Rice Hemoglobins[1]

Gene Cloning, Analysis, and O_2-Binding Kinetics of a Recombinant Protein Synthesized in *Escherichia coli*

Raúl Arredondo-Peter*, Mark S. Hargrove, Gautam Sarath, Jose F. Moran, Joseph Lohrman, John S. Olson, and Robert V. Klucas

Department of Biochemistry, University of Nebraska, The Beadle Center, P.O. Box 880664, Lincoln, Nebraska 68588–0664 (R.A.-P., G.S., J.F.M., J.L., R.V.K.); and Department of Biochemistry and Cell Biology, Rice University, Houston, Texas 77005–1892 (M.S.H., J.S.O.)

Although nonsymbiotic hemoglobins (Hbs) are found in different tissues of dicots and monocots, very little is known about *hb* genes in monocots and the function of Hbs in nonsymbiotic tissues. We report the cloning and analysis of two rice (*Oryza sativa* L.) *hb* genes, *hb1* and *hb2*, that code for plant Hbs. Rice *hb1* and *hb2* genes contain four exons and three introns, as with all of the known plant *hb* genes. At least three copies of the *hb* gene were detected in rice DNA, and analysis of gene expression shows that *hb1* and *hb2* are expressed in leaves but only *hb1* is expressed in roots. A cDNA for rice Hb1 was expressed in *Escherichia coli*, and the recombinant Hb (rHb1) shows an unusually high affinity for O_2 because of a very low dissociation constant. The absorbance spectra of the ferric and deoxyferrous rHb1 indicate that, in contrast to symbiotic Hbs, a distal ligand is coordinated to the ligand-binding site. Mutation of the distal His demonstrates that this residue coordinates the heme Fe of ferric and deoxyferrous rHb1 and stabilizes O_2 in oxy-rHb1. The biochemical properties of rice rHb1 suggest that this protein probably does not function to facilitate the diffusion of O_2.

Hbs are widely distributed throughout higher plants, including both dicots and monocots. Comparison of protein sequences and analysis of gene expression suggest that two families of Hbs, the symbiotic and nonsymbiotic, exist in higher plants (Appleby, 1992; Andersson et al., 1996). Symbiotic Hbs are detected only in N_2-fixing nodules, but not in other plant organs or in non-N_2-fixing plants. The function of Hbs in legume nodules is to facilitate the diffusion of O_2 to the N_2-fixing bacteroids (Appleby, 1984, 1992). Nonsymbiotic Hbs are found in both dicot and monocot plants, and thus appear to be more widely distributed in higher plants than symbiotic Hbs (Appleby, 1992; Andersson et al., 1996).

A Hb that exhibits characteristics of both symbiotic and nonsymbiotic Hbs was isolated from nodules of *Parasponia andersonii*, a nonlegume infected by rhizobia (Appleby et al., 1983). The *P. andersonii hb* gene is expressed in both root nodules and noninfected organs (Bogusz et al., 1988; Appleby, 1992), and codes for a protein that has O_2-binding kinetics similar to symbiotic Hbs, which suggests that both Hbs have a similar function in nodules (Gibson et al., 1989). The *P. andersonii hb* gene contains three introns located identically to the symbiotic *hb* genes (Landsman et al., 1986). A *hb* gene has also been cloned from *Trema tomentosa*, a nonnodulating relative of *P. andersonii* (Bogusz et al., 1988) that is expressed in roots but not in leaves (Bogusz et al., 1988). Nonsymbiotic *hb* genes have been cloned from other dicots such as soybean (*Glycine max* L.) (Andersson et al., 1996) and Arabidopsis (Trevaskis et al., 1997). The soybean *hb* gene is similar to nonsymbiotic *hb*s and it is expressed in diverse organs, with the highest level of expression detected in stems (Andersson et al., 1996).

A Hb transcript from a monocot was cloned by Taylor et al. (1994) using an aleurone cDNA library from barley (*Hordeum vulgare* L.), and its corresponding gene sequence was recently deposited in the GenBank database (accession no. U94968). The amino acid sequence of the predicted barley Hb is similar to nonsymbiotic Hbs. A single copy of the *hb* gene apparently exists in barley. It is expressed in roots of plants grown under normal conditions, but it is expressed at higher levels when the plants are grown under microaerobiosis, suggesting that the expression of the *hb* gene may be associated with the anaerobic response.

[1] This work was supported in part by grants from the National Science Foundation (no. OSR-92552255 to R.V.K.); the U.S. Department of Agriculture-Cooperative State Research Education and Extension Service (no. 95-37305-2441 to R.V.K. and G.S.); the Center for Biotechnology-UNL (to G.S.); the U.S. Public Health Service (nos. GM35649 and HL-47020); and the Robert A. Welch Foundation and the W.M. Keck Foundation (no. C-612 to J.S.O.). R.A.-P. was partially supported by a postdoctoral fellowship from Dirección General de Asuntos del Personal Académico-UNAM and Consejo Nacional de Ciencia y Tecnología, México. J.F.M. is a recipient of a postdoctoral fellowship from the Ministerio de Educación y Cultura, Spain. This work is published as journal series no. 11,902 from the Agricultural Research Division, University of Nebraska.

* Corresponding author; e-mail ra@unlinfo.unl.edu; fax 1–402–472–7842.

Abbreviations: DDBJ, DNA Data Bank of Japan; EST, expressed sequence tags; Hb, hemoglobin; Mb, myoglobin; RGRP, Rice Genome Research Program (Japan); rHb, recombinant hemoglobin.

Incomplete sequences of cDNAs coding for rice Hbs have been deposited in the DDBJ database, and the sequences for the predicted proteins have been published by Andersson et al. (1996). Very little information is available on the *hb* genes in monocots and the biochemical properties of nonsymbiotic Hbs, and nothing is known about the function of these proteins in plants. A nonsymbiotic Hb has been isolated from barley seeds (Duff et al., 1997), and recombinant nonsymbiotic Hbs have been studied (Duff et al., 1997; Trevaskis et al., 1997). These proteins exhibit unusual spectra when reduced and unligated and possess a very high affinity for O_2. In this work we report the cloning and analysis of two *hb* genes, *hb1* and *hb2*, from rice.

A cDNA coding for Hb1 was expressed in *Escherichia coli*, and the recombinant wild-type protein (rHb1) and a mutant in which the distal His was replaced by a Leu were spectroscopically and kinetically characterized. Our results show that rice *hb* genes are similar to other nonsymbiotic *hbs*, and are differentially expressed in roots and leaves of rice plants grown under normal conditions. rHb1 shows spectral characteristics similar to other Hbs, but the spectra for the ferric and deoxyferrous forms are unusual. Furthermore, despite the very high affinity for O_2, rHb1 appears to function through a unique mechanism in which the distal His binds to the heme Fe in the deoxyferrous state, but repositions to stabilize bound O_2, resulting in an extremely low dissociation rate.

MATERIALS AND METHODS

Plant Growth and Total DNA and RNA Isolation

Rice (*Oryza sativa* var. Jackson) seeds were germinated for 5 d and then planted in pots containing vermiculite. Rice plants were grown in a greenhouse at 22°C with light/dark periods of 16 h/8 h and watered with tap water every 3 d and with nutrient solution (Becana et al., 1991) every 6 d. Plants were grown for 5 weeks and then the roots and leaves were collected, washed, and immediately frozen. Total DNA was isolated from roots or leaves using a modification of the cetyltrimethylammonium bromide method (Doyle and Doyle, 1990). Poly(A^+) RNA was isolated from rice roots or leaves using a QuickPrep mRNA purification kit (Pharmacia) and quantitated by spectrophotometry, assuming $1A_{260} = 40\ \mu g/mL$ (Ausubel et al., 1995).

Sequencing of Two cDNAs Coding for Rice Hb1 or Hb2

Rice cDNA clones with sequences similar to plant Hbs were generated by the RGRP (Sasaki et al., 1994) and deposited in the DDBJ database as EST sequences. We obtained the clones C741 and C2576 (DDBJ accession nos. D15507 and D38931, respectively), from the RGRP, which were fully sequenced. Clones C741 and C2576 were named rice Hb1 and Hb2, respectively.

Oligonucleotides and PCR Amplification

Primers were designed for PCR to amplify the *hb1* or *hb2* genes using the sequences that are immediately upstream and downstream from the start and stop codons of the rice Hb1 or Hb2 cDNAs. The oligonucleotide sequences were: Hb1/5′ (sense), 5′-TAAACCAGCTGTCAGGAAGCA-3′; Hb1/3′ (antisense), 5′-AGCAGCT-AGCATGCCTGTCGA-3′; Hb2/5′ (sense), 5′-AGGAATCAAATCGAAGCAGCC-3′; and Hb2/3′ (antisense), 5′-GGAGGTGGAGCAGT-ATATATA-3′. Total rice DNA (approximately 0.5 μg) was used as the template for PCR amplification. PCR components and concentrations were: 0.5 μM of each sense and antisense primer, 200 μM of each dNTP, and 0.4 unit of *Taq* DNA polymerase (Gibco-BRL) in 1× PCR buffer. PCR was done in a final volume of 10 μL using a rapid-cycling apparatus (Idaho Technology, Idaho Falls, ID). Amplification was carried out for 35 cycles at 65°C (for *hb1*) or 60°C (for *hb2*)/30 s for annealing. PCR products were isolated from the agarose gel using the GeneClean kit (Bio 101, Bio-Rad) and then cloned into the vector pCRII (Invitrogen, San Diego, CA) following standard procedures (Sambrook et al., 1989). Cloned fragments were sequenced and DNA sequences were compared with sequences deposited in the GenBank database using the BLAST program (Altschul et al., 1990). Additional computer analyses were done using the GCG package (Genetics Computer Group, Madison, WI).

Southern-Blot Analysis

The clone with the rice *hb1* gene was labeled with digoxigenin-11-dUTP (Boehringer Mannheim) by PCR as described by Lu et al. (1993). Rice total DNA was digested with restriction enzymes, and the samples were Southern-blotted using standard procedures (Sambrook et al., 1989). Membranes were hybridized at 55°C overnight with the rice *hb1* probe, washed at high stringency (60°C) in 2× SSC/0.1% (w/v) SDS twice for 5 min each, and in 0.5× SSC/0.1% (w/v) SDS twice for 15 min each, and then incubated in the NBT/x-Phosphate mix of the Genius kit (Boehringer Mannheim) to develop color.

Expression of *hb* Genes in Rice Organs

Expression of the rice *hb* gene was examined in roots and leaves of 5-week-old plants by RNA-PCR (Wang et al., 1989) using a kit (Cetus). Reverse transcription was done using 30 ng of poly(A^+) RNA as the template and oligo $d(T)_{16}$ as the primer following the manufacturer's protocol. PCR amplification was performed using the specific primers for the amplification of rice Hb1 or Hb2 cDNAs and the same conditions used for the isolation of the rice *hb1* and *hb2* genes (above), but for 40 cycles. For a positive control, ubiquitin transcripts were amplified using primers that were designed from sequences located at positions 109 to 130 (sense: 5′-ATGCAGATCTTCGTGAAGACCC-3′) and 316 to 336 (antisense: 5′-ACCTCCACGAAGGCGCAG-GAC-3′) of a rice ubiquitin cDNA (Nishi et al., 1993). PCR

products were detected in 2% (w/v) agarose gels after staining with EtBr.

Expression of the Rice Hb1 cDNA in *Escherichia coli* and Spectroscopic and Kinetic Characterization of the Recombinant Wild-Type and Mutant Proteins

A cDNA coding for rice Hb1 was amplified by PCR using the primers 5′-*CCATGG*CTCTCGTGGAGGATAAC-3′ (sense) and 5′-*GAATTC*TCACTCCGCGGGCTTCATCTC-3′ (antisense) that are located at the start and stop codons of the *hb1* gene, respectively, and were degenerated with restriction sequences for *Nco*I or *Eco*RI (underlined). PCR conditions were the same as above. The PCR fragment for Hb1 was subcloned into the expression vector pET28a (Novagen, Madison, WI) and transformed in *E. coli* as described by Hargrove et al. (1997) and Arredondo-Peter et al. (1997). The Kunkel method (Kunkel, 1985) of mutagenesis was used to generate an H74L (His→Leu) mutant starting from the cloned cDNA for Hb1. Recombinant rice Hbs (wild type and mutant) were purified by $(NH_4)_2SO_4$ precipitation and chromatography on DEAE-cellulose, and then partially sequenced from their N terminus using standard procedures (Jun et al., 1994a, 1994b). Pure recombinant Hbs were spectrophotometrically characterized as described by Arredondo-Peter et al. (1997), and O_2- and CO-binding affinities were determined as described by Hargrove et al. (1997).

RESULTS AND DISCUSSION

Sequencing of Two Rice cDNAs That Code for Hb1 and Hb2 and Analysis of the Predicted Proteins

Partial sequences (approximately 300 bp) of two cDNAs, corresponding to the clones C741 and C2576, with high similarity to nonsymbiotic *hb* genes, have been generated by the RGRP program and deposited in the DDBJ database (Sasaki et al., 1994). We fully sequenced the clones C741 and C2576 in both directions, and the sequences were compared with sequences deposited in the GenBank database. Clones C741 and C2576 were found to be highly similar to plant Hbs, so they were designated Hb1 and Hb2, respectively. The rice Hb1 and Hb2 clones are 812 and 786 bp in length, contain putative polyadenylation signals at positions 772 and 747, and code for predicted proteins of 166 and 169 amino acid residues, respectively (Fig. 1).

Sequence comparisons show that the predicted Hb1 and Hb2 proteins are 93% similar to each other, and that the rice Hbs are 68 to 82% similar to nonsymbiotic Hbs and about 50% similar to symbiotic Hbs. Rice Hbs contain distal (H77) and proximal (H112) His residues, as well as the P52, F58, F82, and F122 that are conserved in plant Hbs (Fig. 2) (Arredondo-Peter and Escamilla, 1991). Rice Hb1 and Hb2 also contain a single Cys residue, C86, that is highly conserved in nonlegume Hbs (Arredondo-Peter and Escamilla, 1991; Taylor et al., 1994; Andersson et al., 1996). Andersson et al. (1996) reported a second Cys at position 92 (numbering as in Fig. 2) from partial sequences of rice Hb1 and Hb2. However, after sequencing many clones of rice Hbs we did

A

```
CCACGCGTCCGGTTGTTTTCAGAGCCCAGCTAGCTCTCGATCATTTGTTA   50
CAGAGAAATTGATCAAAGCAGGAAATTAAACCAGCTGTCAGGAAGCAATG  100
                                                 M    1

GCTCTCGTGGAGGATAACAATGCCGTAGCGGTGAGCTTCAGCGAGGAGCA  150
 A  L  V  E  D  N  N  A  V  A  V  S  F  S  E  E  Q   18

GGAGGCGCTGGTGCTCAAGTCATGGGCGATCTTGAAGAAGGATTCCGCCA  200
  E  A  L  V  L  K  S  W  A  I  L  K  K  D  S  A    34

ATATTGCCCTCCGCTTCTTCTTGAAGATCTTCGAGGTCGCGCCGTCGGCG  250
N  I  A  L  R  F  F  L  K  I  F  E  V  A  P  S  A   51

AGCCAGATGTTCTCGTTCCTGCGAAACTCCGACGTGCCGCTCGAGAAGAA  300
 S  Q  M  F  S  F  L  R  N  S  D  V  P  L  E  K  N   68

CCCCAAGCTCAAGACCCACGCCATGTCCGTCTTCGTCATGACATGCGAGG  350
  P  K  L  K  T  H  A  M  S  V  F  V  M  T  C  E     84

CCGCCGCGCAGCTGCGGAAAGCCGGGAAGGTCACCGTGAGAGACACCACC  400
A  A  A  Q  L  R  K  A  G  K  V  T  V  R  D  T  T   101

CTCAAGAGGCTCGGCGCCACGCACCTCAAGTACGGCGTCGGAGACGCCCA  450
 L  K  R  L  G  A  T  H  L  K  Y  G  V  G  D  A  H  118

CTTCGAGGTGGTGAAGTTCGCGCTGCTTGACACGATCAAGGAGGAGGTTC  500
  F  E  V  V  K  F  A  L  L  D  T  I  K  E  E  V    134

CGGCGGACATGTGGAGCCCGGCGATGAAGAGCGCGTGGAGCGAAGCCTAC  550
P  A  D  M  W  S  P  A  M  K  S  A  W  S  E  A  Y   151

GACCACCTGGTCGCTGCCATCAAGCAGGAGATGAAGCCCGCGGAGTGATC  600
 D  H  L  V  A  A  I  K  Q  E  M  K  P  A  E  *     166

GACAGGCATGCTAGCTGCTCCACCTCCATGATCCTCGCCTCGCGAGTCCG  650
ATTAGCTTTTGTTGCTTTCAAATGCTCGTTTCATATTCATCGTGTCCCAC  700
AAAAAAAGGAGTGTGTATGTGGTGTACGATTGGTGCACGCTCCTGCTGTT  750
TTCTTCTCGTGATAAGACATAAATAAAGATGGTTTTCTACGCTAAAAAAA  800
AAAAAAAAAAAG                                        812
```

B

```
CCACGCGTCCGGTTGAGTTGAATTGAGCTCGAATTGTACTCGATTCACCA   50
CACAGAGGAATCAAATCGAAGCAGCCATGGCTCTCGTGGAGGGAAACAAC  100
                        M  A  L  V  E  G  N  N      8

GGCGTGTCGGGGGGAGCGGTCAGCTTCAGCGAGGAGCAGGAGGCGCTTGT  150
 G  V  S  G  G  A  V  S  F  S  E  E  Q  E  A  L  V   25

GCTCAAGTCGTGGGCCATCATGAAGAAGGATTCCGCCAACATTGGACTCC  200
  L  K  S  W  A  I  M  K  K  D  S  A  N  I  G  L     41

GCTTCTTCTTGAAGATCTTCGAGGTCGCGCCGTCGGCGAGCCAGATGTTC  250
R  F  F  L  K  I  F  E  V  A  P  S  A  S  Q  M  F   58

TCGTTCCTGCGCAACTCCGACGTGCCGCTCGAGAAGAACCCCAAGCTCAA  300
 S  F  L  R  N  S  D  V  P  L  E  K  N  P  K  L  K   75

GACCCACGCCATGTCCGTCTTCGTCATGACATGTGAGGCCGCCGCGCAGC  350
  T  H  A  M  S  V  F  V  M  T  C  E  A  A  A  Q     91

TGCGGAAAGCCGGGAAGGTCACCGTGAGAGACACCACCCTGAAGAGGCTC  400
L  R  K  A  G  K  V  T  V  R  D  T  T  L  K  R  L   108

GGCGCCACGCACTTCAAGTACGGCGTCGGAGACGCCCACTTTGAGGTGAC  450
 G  A  T  H  F  K  Y  G  V  G  D  A  H  F  E  V  T  125

GAGGTTCGCGCTGCTTGAGACGATCAAGGAGGCGGTTCCGGTGGACATGT  500
  R  F  A  L  L  E  T  I  K  E  A  V  P  V  D  M    141

GGAGCCCCGCGATGAAGAGCGCGTGGAGCGAAGCCTACAACCAACTGGTC  550
W  S  P  A  M  K  S  A  W  S  E  A  Y  N  Q  L  V   158

GCGGCCATCAAGCAGGAGATGAAGCCTGCTGAGTGATATATATACTGCTC  600
 A  A  I  K  Q  E  M  K  P  A  E  *                 169

CACCTCCATGATCCTCGCTGATCAACTTTGTTGCATTGTGCTCGTTCAAT  650
ATTCCTCGCCCCACAAAAGGGACTTTTGTCGGTGTGTGTATGTCCAAATG  700
ATTAAATCAACTGCTGTTTTGTTCTATGTAAGATACATAACTCATAAATA  750
AAGATGGTTTTCTACATGCAAAAAAAAAAAAAAAG                 786
```

Figure 1. Nucleotide and deduced protein sequences of clones C741 (A) and C2576 (B) that code for rice Hb1 and Hb2, respectively. Sequences used to design the sense and antisense primers for the specific amplification of Hb1 and Hb2 are underlined. Putative polyadenylation signals are double-underlined.

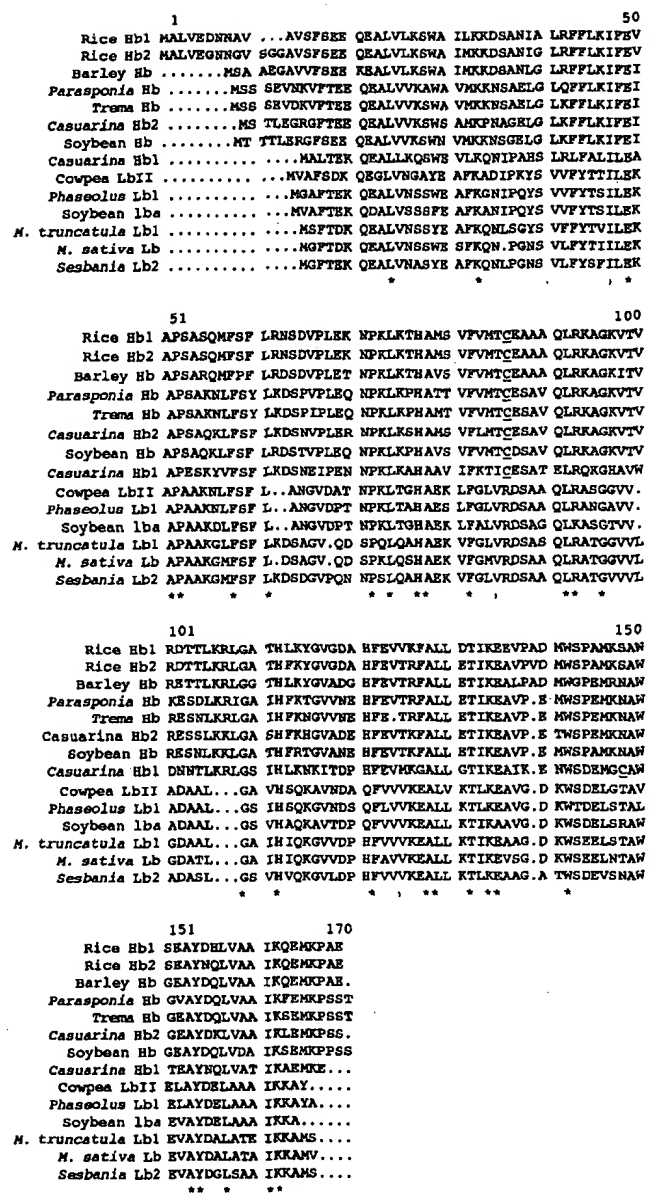

Figure 2. Sequence alignment of nonsymbiotic and selected symbiotic Hbs. Distal (H77) and proximal (H112) His residues are shown in bold type, and Cys residues are underlined. Intron I is at position 47, intron II is at position 85, and intron III is at position 125; asterisks show the most conserved residues. Amino acid sequences were obtained from the GenBank database using the following accession numbers: U01228 (barley Hb), M36509 (*Parasponia andersonii* Hb), Y00296 (*Trema tomentosa* Hb), X53950 (*Casuarina glauca* Hb2), U47143 (soybean Hb), L28826 (*C. glauca* Hb1), U33207 (cowpea LbII), K03152 (*Phaseolus vulgaris* Lb1), V00453 (soybean Lba), X57732 (*Medicago truncatula* Lb1), M32883 (*M. sativa* Lb), and X13815 (*Sesbania rostrata* Lb2). Alignment of sequences was done using the PileUp routine of the GCG program.

not detect a Cys C92, but rather a Leu L92, which is highly conserved in plant Hbs (Fig. 2). Thus, our results support the conclusion that rice Hbs contain only one Cys residue, which is located at position 86.

Cloning and Analysis of Rice *hb* Genes

To clone the gene for rice Hb, we used total rice DNA as the template with specific primers for rice Hb1 or Hb2, which generated PCR products of approximately 900 bp. The PCR products were purified, cloned, and sequenced (Fig. 3), and the coding sequences of the PCR products were identical to cDNA sequences for Hb1 and Hb2; therefore, they were the rice *hb1* and *hb2* genes.

The *hb* gene and cDNA sequences were compared to identify the exon and intron (IVS) sequences of the rice *hb* genes. Rice *hb* genes have four exons and three introns, with the introns located at the same position as all of the known plant *hb* genes. The exon/intron boundaries of the rice *hb* genes are identical to the boundaries of the *Parasponia andersonii hb* gene (Appleby et al., 1988) (Fig. 3). Homologous introns of rice *hb1* and *hb2* were very conserved, with 92% similarity for IVS-II, 84% similarity for IVS-III, and 72% similarity for IVS-I. The highly conserved location of the introns and sequences of the exon/intron boundaries in monocot and dicot *hb* genes suggest that the ancestral *hb* gene of flowering plants had three introns in an identical location.

When rice DNA was digested with restriction enzymes and then subjected to Southern blotting with the rice *hb1* probe at high stringency, between one and three hybridizing fragments were detected (Fig. 4). Rice *hb* genes have no restriction sites for the enzymes used to cut the DNA, so at least three copies of the *hb* gene exist in rice. The existence of cDNAs encoding for Hb1 and Hb2 (Fig. 1) indicates that *hb1* and *hb2* are functional genes in rice.

Expression of *hb* Genes in Rice Organs

In contrast to symbiotic *hb* genes, which are expressed only in nodules of N_2-fixing plants, nonsymbiotic Hb transcripts have been reported to exist in many tissues, including: (a) root meristems of *Trema tomentosa* (Bogusz et al., 1988), (b) root vascular bundles of transgenic tobacco (*Nicotiana tabacum*) (Bogusz et al., 1990), (c) seed aleurone and roots grown under microaerobiosis of barley (*Hordeum vulgare*) (Taylor et al., 1994), and (d) diverse organs of soybean (*Glycine max*) (Andersson et al., 1996) and Arabidopsis (Trevaskis et al., 1997).

To determine the pattern of *hb* gene expression in rice, we isolated poly(A^+) RNA from roots and leaves, and then subjected it to PCR using specific primers for rice Hb1 or Hb2. Amplification products were detected for Hb1 in roots, and for Hb1 and Hb2 in leaves (Fig. 5). Hb transcripts of approximately 550 bp were cloned and sequenced, and the resulting sequences were identical to those of the rice Hb1 and Hb2 cDNAs, indicating that *hb* genes are detectable and functional in rice roots and leaves. We did not detect any Hb2 transcripts in rice roots using our protocols, suggesting that the *hb2* gene is probably not expressed in the roots of rice grown under normal growth conditions. The pattern of expression of rice *hb* genes is similar to the expression of *hb* genes in Arabidopsis reported by Trevaskis et al. (1997). The differential expression of *hb1* and *hb2* genes in the rice plant indicates that these genes

```
hb1    1 ATGGCTCTCGTGGAGGATAACAATGCCGT.........AGCGGTGAGCTT 41
hb2    1 ATGGCTCTCGTGGAGGGAAACAACGGCGTGTCGGGGGGAGCGGTCAGCTT 50

      42 CAGCGAGGAGCAGGAGGCGCTGGTGCTCAAGTCATGGGCGATCTTGAAGA 91
      51 CAGCGAGGAGCAGGAGGCGCTTGTGCTCAAGTCGTGGGCCATCATGAAGA 100

      92 AGGATTCCGCCAATATTGCCCTCCGCTTCTTCTTGAAgtatgtac..atg 139
     101 AGGATTCCGCCAACATTGGACTCCGCTTCTTCTTGAAgtatgtactacta 150

     140 cgtgttacta.........ccattt........ctctttttgcggaatc 171
     151 catgctactactcactagtgccatttttcgcagagtcagtttacggcacc 200

     172 agagattgggtt.tgtgaagcat..taaattgagcaatgcat..ttcgct 216
     201 aaccattggtttctgtagtgcatcatacattttggtgtgcgtgatttggt 250

     217 gatacatgtgtgtctgattgtgttgtagGATCTTCGAGGTCGCGCCGTCG 266
     251 gctac.......tatatttttattccagGATCTTCGAGGTCGCGCCGTCG 293

     267 GCGAGCCAGATGTTCTCGTTCCTGCGAAACTCCGACGTGCCGCTCGAGAA 316
     294 GCGAGCCAGATGTTCTCGTTCCTGCGCAACTCCGACGTGCCGCTCGAGAA 343

     317 GAACCCCAAGCTCAAGACCCACGCCATGTCCGTCTTCGTCATGgtaatac 366
     344 GAACCCCAAGCTCAAGACCCACGCCATGTCCGTCTTCGTCATGgtaatac 393

     367 taccatcattat...ttcaggca..agtaaatttgttgtggtagtagaca 411
     394 taccatcagtattgctgcagacaagagtaaatttgttgtagtagtaggca 443

     412 ctgacagaatgtgtgcgtgcgtcgcgatcaatcgatattgcagACATGCG 461
     444 ctgacagaatgtgtgcgtgtgtcgtgatcaatcgatgttgcagACATGTG 493

     462 AGGCCGCCGCGCAGCTGCGGAAAGCCGGGAAGGTCACCGTGAGAGACACC 511
     494 AGGCCGCCGCGCAGCTGCGGAAAGCCGGGAAGGTCACCGTGAGAGACACC 543

     512 ACCCTCAAGAGGCTCGGCGCCACGCACCTCAAGTACGGCGTCGGAGACGC 561
     544 ACCCTGAAGAGGCTCGGCGCCACGCACTTCAAGTACGGCGTCGGAGACGC 593

     562 CCACTTCGAGgtacagtgatccccaatggctgcctgcgctccattcgatc 611
     594 CCACTTTGAGgtagagtgatatccagttgtt......gctccattcgatc 637

     612 gacatgaaac....ttgatcgttttctgatcgtgtctttgtcgaacaacg 657
     638 gacatggaactctatcgatcgtgttctgattgtggccttgtcgaa..... 682

     658 tacatgcgatcgwtcgatcgtgtawacagGTGGTGAAGTTCGCGCTGCTT 707
     683 ..........cgttc......gtaaacagGTGACGAGGTTCGCGCTGCTT 716

     708 GACACGATCAAGGAGGAGGTTCCGGCGGACATGTGGAGCCCGGCGATGAA 757
     717 GAGACGATCAAGGAGGCGGTTCCGGTGGACATGTGGAGCCCCGCGATGAA 766

     758 GAGCGCGTGGAGCGAAGCCTACGACCACCTGGTCGCTGCCATCAAGCAGG 807
     767 GAGCGCGTGGAGCGAAGCCTACAACCAACTGGTCGCGGCCATCAAGCAGG 816

     808 AGATGAAGCCCGCGGAGTGA 827
     817 AGATGAAGCCTGCTGAGTGA 836
```

Figure 3. Sequence alignment of rice *hb1* and *hb2* genes. Coding and noncoding sequences are shown as upper or lowercase, respectively. Sequences flanking the exon/intron boundaries that are conserved in rice and *P. andersonii* sp. (Appleby et al., 1988) *hb* genes are underlined.

are not linked to each other, and that each *hb* gene is probably regulated by different promoter sequences and *trans*-acting factors.

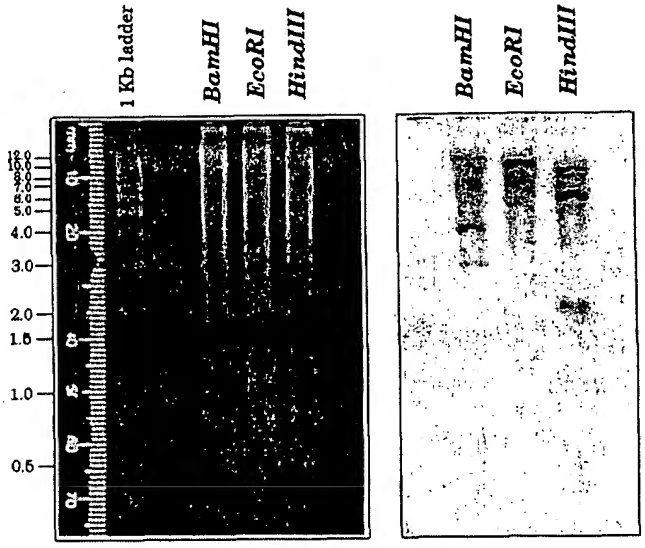

Figure 4. Southern-blot analysis of total rice DNA using the rice *hb1* gene as a probe. Signals were detected by colorimetry using the chromogenic mix NBT/x-phosphate. Molecular size markers are shown in kilobars.

Expression of the Rice Hb1 cDNA in *E. coli* and Spectroscopic and Kinetic Characterization of the Recombinant Wild-Type and Mutant Protein

Nonsymbiotic Hbs are of low abundance in plant tissues and thus are difficult to isolate and purify from plants. The synthesis of recombinant proteins provides a useful method of producing large amounts of protein for bio-

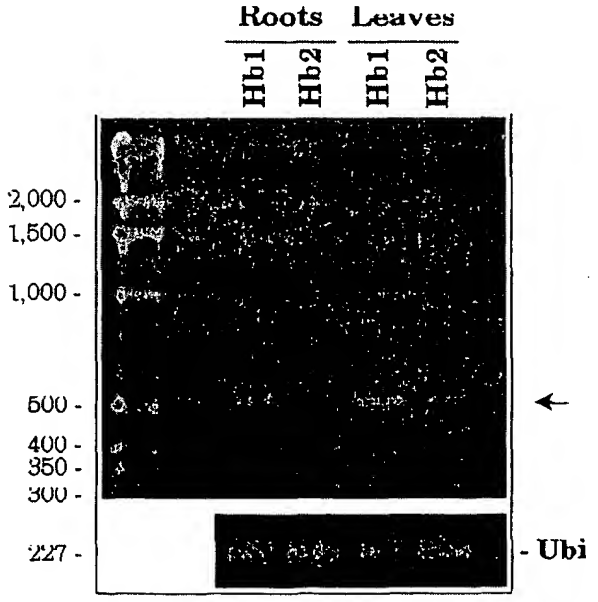

Figure 5. Agarose gel electrophoresis of the RNA-PCR products that were obtained using as a template poly(A$^+$) RNA isolated from rice roots or leaves and specific primers for rice Hb1 or Hb2. Arrow shows the approximately 550-bp fragments that were cloned and sequenced. Rice ubiquitin (Ubi) was used as positive control. Molecular size markers are shown in base pairs.

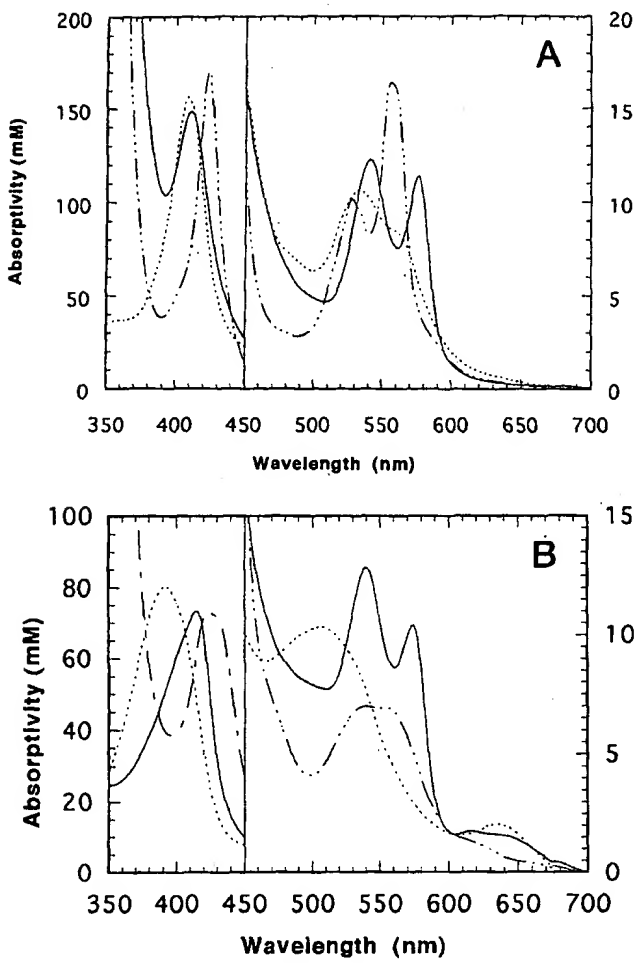

Figure 6. Absorption spectra of rice wild-type rHb1 (A) and the H77L mutant of rHb1 (B). Dashed lines, Ferric; solid lines, oxygenated; and combination dashed/solid lines, ferrous forms of Hb.

chemical studies. We prepared a recombinant Hb by subcloning a cDNA for rice Hb1 into the vector pET28a and then expressing it in *E. coli*. The N terminus of the rHb1 was determined to be ALVEDNNAVAV, which is identical to the predicted sequence of the Hb1 cDNA (Fig. 1A), indicating that the correct recombinant protein had been synthesized.

Analysis of the rHb1 shows that it exhibits spectral characteristics that are similar to Hbs (Fig. 6A), including reversible binding of O_2. However, rHb1 exhibits some distinctive absorption bands both in the ferric and the deoxyferrous forms. Differences in the globin-heme linkage were apparent from the deoxyferrous spectrum because, at pH 7.0, the unligated ferrous state exhibits two peaks at 526 and 556 nm, which is similar to the absorption spectra of Cyt *b*, where the heme is hexacoordinate and the Fe is principally in the low-spin state (Smith, 1978; Weiss and Ziganke, 1978). These spectra are in marked contrast to the symbiotic plant Hbs and animal Mbs and Hbs, in which the Fe is pentacoordinate and displays a broad peak centered at 556 nm in their deoxyferrous form (Appleby, 1974, 1992). Thus, it is apparent that the ferric and deoxyferrous forms of rHb1 contain a distal ligand that was identified as His74. Absorbance spectra of the ferric, oxy-, and deoxyferrous forms of an H74L mutant of rHb1 shows no evidence of His coordination (Fig. 6B), and the ferric peak at 405 nm suggests that the ligand-binding site is partially occupied by a water molecule. Furthermore, the addition of exogenous imidazole to ferric or deoxyferrous H74L mutant results in a spectrum identical to the corresponding form of the wild-type rHb1.

Kinetic analysis of ligand binding shows that the rHb1 has an unusually high affinity for O_2 (Table I). The O_2-association constant of rHb1 is similar to other O_2 storage and transport proteins, such as soybean Lb*a*, however, the high affinity of rHb1 for O_2 results from a very low dissociation constant. Similar values for the dissociation constant have been reported for barley (Duff et al., 1997) and Arabidopsis (Trevaskis et al., 1997) recombinant Hbs, and suggest that these proteins are not involved in O_2 transport through facilitated diffusion. Rate constants of rice rHb1 for the reaction with CO are similar to O_2-transport proteins, which indicates that the unique O_2 reactivity is a result of specific interactions between O_2 and the protein.

Removal of His74 had a profound effect on the rate constants for O_2 binding. The association constant increases approximately 9-fold as a result of the H74L mutation, and the dissociation constant increases nearly 1000 times. This is similar to the effect of removal of the distal His in Mb (Springer and Sligar, 1987), but is different from the smaller effect seen with the same mutation in soybean Lb*a* (Hargrove et al., 1997) (Table I). These results suggest that His74 forms a strong H^+ bond with bound O_2, which contributes to the slow dissociation constant. However, the rate constant for the mutant, which is similar to soybean Lb, combined with the extremely low value for the wild-

Table I. *Rate and equilibrium constants for the reaction of O_2 and CO with rice wild-type rHb1 and the H77L mutant of rHb1*

Protein	k'_{O_2}[a]	k_{O_2}[b]	K_{O_2}[c]	k'_{CO}[a]	k_{CO}[b]	K_{CO}[c]
	$\mu M^{-1}\ s^{-1}$	s^{-1}	μM^{-1}	$\mu M^{-1}\ s^{-1}$	s^{-1}	μM^{-1}
Rice rHb1	68	0.038	1800	7.2	0.001	7,200
rHb1-H77L	620	51	12	150	0.002	75,000
Soybean Lb*a*[d]	130	5.6	23	16	0.0084	1,900
Lb*a*-H61L[d]	400	24	16	170	0.0024	71,000

[a] Association constant. [b] Dissociation constant. [c] Ligand affinity. [d] Values from Hargrove et al. (1997).

type rHb1, suggests that there might be an additional mechanism limiting O_2 dissociation.

Nonsymbiotic Hb from rice, and presumably all Hbs of this class, appear to operate through a unique mechanism. His H74 coordinates the heme Fe in the deoxy-protein, but when O_2 binds the side chain moves to a position from which it can form a stabilizing interaction with bound O_2. This is very different from the behavior of Mb mutants, in which a His coordinates the heme Fe. These proteins are very unstable and do not react with O_2 due to an extraordinarily high autooxidation rate (Duo et al., 1995).

Function of Nonsymbiotic Hbs

The function of plant Hbs in nonsymbiotic tissues is not known; however, it has been hypothesized that the role of nonsymbiotic Hbs is probably not to facilitate the diffusion of O_2 in roots, but rather to sense levels of O_2 (Appleby et al., 1988). Appleby et al. (1988) suggested that under normal aerobic conditions Hb would be oxygenated, and that under O_2-limiting conditions, deoxyferrous Hb levels would increase and trigger an anaerobic response. Although O_2-sensor heme proteins have been described in other systems (Gilles-Gonzalez et al., 1991, 1994, 1995; Gilles-Gonzalez and Gonzalez, 1993), the above hypothesis was questioned recently by Andersson et al. (1996), who suggest that Hbs may function as O_2 carriers in metabolically active tissues.

Our results with rice plants, including the expression profiles in aerobic tissues (Fig. 5), expected levels of Hbs in tissues, and kinetic parameters of Hb1 (Table I), suggest that these proteins are unlikely to facilitate the diffusion of O_2. Also, because of the extremely low dissociation constant exhibited by rHb1, it is unlikely that rice Hbs function as an effective O_2 donor to other proteins, such as mitochondrial oxidases. However, under specific circumstances, such as in barley roots grown under microaerobic conditions, in which high levels of Hb transcripts have been detected (Taylor et al., 1994), nonsymbiotic Hbs may have other functions, including participation in the anaerobic response and possibly in specific metabolic aspects of dedifferentiated tissues. Recently, a multitude of functions have been suggested for nonplant Hbs, which includes the transport of ligands other than O_2 (such as NO or CO), interaction with small organic molecules, O_2 scavenging, or formation of complexes with regulatory proteins (Götz et al., 1994; Giardina et al., 1995; Goldberg, 1995; Jia et al., 1996). The pattern of expression of nonsymbiotic *hb* genes in plants and the biochemical properties of the recombinant Hbs suggest that these proteins have other functions besides O_2 transport, which are yet to be determined.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to the RGRP for the kind gift of the rice clones C741 and C2576, and to Dr. H.E. Bockelman (U.S. Department of Agriculture-Agricultural Research Service) for supplying the rice seeds.

Received May 13, 1997; accepted July 30, 1997.

Sequences reported in this work were deposited in the GenBank database under the accession nos. U76028, U76029, U76030, and U76031 for the *hb1* and *hb2* genes and Hb1 and Hb2 cDNA, respectively.

LITERATURE CITED

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403–410

Andersson CR, Ostergaard E, Llewellyn DJ, Dennis ES, Peacock WJ (1996) A new hemoglobin gene from soybean: a role for hemoglobin in all plants. Proc Natl Acad Sci USA **93:** 5682–5687

Appleby CA (1974) Leghemoglobin. *In* A Quispel, ed, The Biology of Nitrogen Fixation. American Elsevier Publishing Co., New York, pp 521–554

Appleby CA (1984) Leghemoglobin and *Rhizobium* respiration. Annu Rev Plant Physiol **35:** 443–478

Appleby CA (1992) The origin and functions of haemoglobin in plants. Sci Progress **76:** 365–398

Appleby CA, Bogusz D, Dennis ES, Peacock WJ (1988) A role for hemoglobin in all plant roots? Plant Cell Environ **11:** 359–367

Appleby CA, Tjepkema JD, Trinick MJ (1983) Hemoglobin in a nonleguminous plant *Parasponia*: possible genetic origin and function in nitrogen fixation. Science **220:** 951–953

Arredondo-Peter R, Escamilla E (1991) A consensus sequence of plant hemoglobins. Plant Mol Biol Reporter **9:** 195–207

Arredondo-Peter R, Moran JF, Sarath G, Luan P, Klucas RV (1997) Molecular cloning of the cowpea (*Vigna unguiculata*) leghemoglobin II gene and expression of its cDNA in *Escherichia coli*. Purification and characterization of the recombinant protein. Plant Physiol **114:** 493–500

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1995) Current Protocols in Molecular Biology. Wiley, New York

Becana M, Salin ML, Ji L, Klucas RV (1991) Flavin-mediated reduction of ferric leghemoglobin from soybean nodules. Planta **183:** 575–583

Bogusz D, Appleby CA, Landsmann J, Dennis ES, Trinick MJ, Peacock WJ (1988) Functioning of haemoglobin genes in a non-nodulating plant. Nature **331:** 178–180

Bogusz D, Llewellyn DJ, Craig S, Dennis ES, Appleby CA, Peacock WJ (1990) Nonlegume hemoglobin genes retain organ-specific expression in heterologous transgenic plants. Plant Cell **2:** 633–641

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus **12:** 13–15

Duff SMG, Wittenberg JB, Hill RD (1997) Expression, purification and properties of recombinant barley (*Hordeum* spp.) hemoglobin: optical spectra and reactions with gaseous ligands. J Biol Chem **272:** 16746–16752

Duo Y, Admiraal SJ, Ikeda-Saito M, Krzywda S, Wilkinson AJ, Li T, Olson JS, Prince RC, Pickering IJ, George GN (1995) Alteration of axial coordination by protein engineering in myoglobin. Bisimidazole ligation in the His64 → Val/Val68 → His double mutant. J Biol Chem **270:** 15993–16001

Giardina B, Messana I, Scatena R, Castagnola M (1995) The multiple functions of hemoglobin. Crit Rev Biochem Mol Biol **30:** 165–196

Gibson QH, Wittenberg JB, Wittenberg BA, Bogusz D, Appleby CA (1989) The kinetics of ligand binding to plant hemoglobins. J Biol Chem **264:** 100–107

Gilles-Gonzalez MA, Ditta GS, Helinski DR (1991) A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*. Nature **350:** 170–172

Gilles-Gonzalez MA, Gonzalez G (1993) Regulation of the kinase activity of heme protein FixL from the two-component system FixL/FixJ of *Rhizobium meliloti*. J Biol Chem **268:** 16293–16297

Gilles-Gonzalez MA, Gonzalez G, Perutz MF (1994) Heme-based sensors, exemplified by the kinase FixL, are a new class of heme protein with distinctive ligand binding and autoxidation. Biochemistry **33:** 8067–8073

Gilles-Gonzalez MA, Gonzalez G, Perutz MF (1995) Kinase activity of oxygen sensor FixL depends on the spin state of its heme iron. Biochemistry **34:** 232–236

Goldberg DE (1995) The enigmatic oxygen-avid hemoglobin of *Ascaris*. BioEssays **17:** 177–182

Götz FM, Hertel M, Groschel-Stewart U (1994) Fatty acid binding of myoglobin depends on its oxygenation. Biol Chem Hoppe-Seyler **375:** 387–392

Hargrove MS, Barry JK, Brucker EA, Berry MB, Phillips GN, Olson JS, Arredondo-Peter R, Dean JM, Klucas RV, Sarath G (1997) Characterization of recombinant soybean leghemoglobin *a* and apolar distal histidine mutants. J Mol Biol **267:** 1032–1042

Jia L, Bonaventura C, Bonaventura J, Stamler JS (1996) S-nitrosohaemoglobin: a dynamic activity of blood involved in vascular control. Science **380:** 221–226

Jun H, Sarath G, Moran JF, Becana M, Klucas RV, Wagner FW (1994a) Characteristics of modified leghemoglobins isolated from soybean (*Glycine max* Merr.) Root Nodules. Plant Physiol **104:** 1231–1236

Jun H, Sarath G, Wagner FW (1994b) Detection and purification of modified leghemoglobins from soybean root nodules. Plant Sci **100:** 31–40

Kunkel TA (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc Natl Acad Sci USA **82:** 488–492

Landsman J, Dennis ES, Higgins TJV, Appleby CA, Kortt AA, Peacock WJ (1986) Common evolutionary origin of legume and non-legume plant haemoglobins. Nature **324:** 166–168

Lu YH, Negre S, Leroy P, Bernard M (1993) PCR-mediated screening and labeling of DNA from clones. Plant Mol Biol Reporter **11:** 345–349

Nishi R, Hashimoto H, Kidou S, Uchimiya H, Kato A (1993) Isolation and characterization of a rice cDNA which encodes a ubiquitin protein and a 52 amino acid extension protein. Plant Mol Biol **22:** 159–161

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sasaki T, Song J, Koga-Ban Y, Matsui E, Fang F, Higo H, Nagasaki H, Hori M, Miya M, Murayama-Kayano E, and others (1994) Toward cataloguing all rice genes: large scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. Plant J **6:** 615–624

Smith L (1978) Bacterial cytochromes and their spectral characterization. Meth Enzymol **53:** 202–212

Springer BA, Sligar SG (1987) High-level expression of sperm whale myoglobin in *Escherichia coli*. Proc Natl Acad Sci USA **84:** 8961–8965

Taylor ER, Nie XZ, MacGregor AW, Hill RD (1994) A cereal haemoglobin gene is expressed in seed and root tissues under anaerobic conditions. Plant Mol Biol **24:** 853–862

Trevaskis B, Watts R, Andersson CR, Llewellyn DJ, Hargrove MS, Olson JS, Dennis ES, Peacock WJ (1997) Two hemoglobin genes in *Arabidopsis thaliana*: a new family of plant hemoglobins that may be ancestral to the leghemoglobins. Proc Natl Acad Sci USA (in press)

Wang AM, Doyle MV, Mark DF (1989) Quantitation of mRNA by the polymerase chain reaction. Proc Natl Acad Sci USA **86:** 9717–9721

Weiss H, Ziganke B (1978) Purification of cytochrome *b* from *Neurospora crassa* and other sources. Methods Enzymol **53:** 212–221

Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 12230–12234, October 1997
Plant Biology

Two hemoglobin genes in *Arabidopsis thaliana*: The evolutionary origins of leghemoglobins

BEN TREVASKIS*†‡, RICHARD. A. WATTS†‡, CAROL R. ANDERSSON†§, DANNY J. LLEWELLYN†, MARK S. HARGROVE¶, JOHN S. OLSON¶, ELIZABETH S. DENNIS†‖, AND W. JAMES PEACOCK†

*Cooperative Research Centre for Plant Cell Biology, Australian National University, Australian Capital Territory, 2601 Australia; †Division of Plant Industry, Commonwealth Scientific and Industrial Research Organisation, GPO Box 1600, Canberra, Australian Capital Territory, 2601 Australia; ‡Department of Biochemistry and Molecular Biology, Australian National University, Australian Capital Territory, 2601 Australia; and ¶Department of Biochemistry and Cell Biology, Rice University, 6100 Main Street, Houston, TX 77005-1892

Communicated by Marshall D. Hatch, Commonwealth Scientific and Industrial Research, Canberra, Australia, August 27, 1997 (received for review June 27, 1997)

ABSTRACT **We cloned two hemoglobin genes from *Arabidopsis thaliana*. One gene, *AHB1*, is related in sequence to the family of nonsymbiotic hemoglobin genes previously identified in a number of plant species (class 1). The second hemoglobin gene, *AHB2*, represents a class of nonsymbiotic hemoglobin (class 2) related in sequence to the symbiotic hemoglobin genes of legumes and *Casuarina*. The properties of these two hemoglobins suggest that the two families of nonsymbiotic hemoglobins may differ in function from each other and from the symbiotic hemoglobins. *AHB1* is induced, in both roots and rosette leaves, by low oxygen levels. Recombinant AHB1 has an oxygen affinity so high as to make it unlikely to function as an oxygen transporter. *AHB2* is expressed at a low level in rosette leaves and is low temperature-inducible. AHB2 protein has a lower affinity for oxygen than AHB1 but is similar to AHB1 in having an unusually low, pH-sensitive oxygen off-rate.**

Leghemoglobins, the first hemoglobins to be identified in higher plants (1), are found in the nodules of legumes where they are believed to transport oxygen to nitrogen-fixing endosymbiotic bacteria (2). Nodule-expressed hemoglobins also have been identified in *Casuarina glauca*, a nitrogen-fixing nonlegume (3). Collectively, these nodule-expressed hemoglobins are called "symbiotic hemoglobins" (4). In addition to symbiotic hemoglobins, these plants have another hemoglobin that shows only limited amino acid sequence similarity to the symbiotic hemoglobins. Genes encoding this type of hemoglobin also have been cloned from the nitrogen-fixing elm *Parasponia andersonii* (5, 6) and from plants that do not fix nitrogen, including monocots (6–8). These "nonsymbiotic" hemoglobins are typically expressed at low levels in root tissue. The *Parasponia* hemoglobin is unusual in that it appears to be bi-functional; it transports oxygen in nitrogen-fixing nodules and, like the nonsymbiotic hemoglobins, is expressed in root tissue (6). The function of nonsymbiotic hemoglobins in normal plant tissues has not been established. To examine the function of nonsymbiotic hemoglobins, we have cloned hemoglobins from *Arabidopsis thaliana*.

MATERIALS AND METHODS

Plant Growth and Stress Treatments. Arabidopsis C24 ecotype seeds were surface sterilized, then grown on MS agar with 3% sucrose, 40–50 plants/plate, at 22°C (16/8 h light/dark cycle) until harvest. Plants used for *AHB1* induction analysis were grown on agar without sucrose. Dehydration, chilling, and hypoxia treatments were performed as outlined in Dolferus *et al.* (9). All treatments, except chilling (24 h), were for 12 h unless specified. Sucrose treatment involved placing plantlets in MS with 1% sucrose overnight. RNA was extracted using the method of Logemann *et al.* (10) and separated on formaldehyde gels as described by Dolferus *et al.* (9); RNA gel blotting, hybridization, and posthybridization washes also were done as described by Dolferus *et al.* RNA loadings were calibrated with an *Arabidopsis* Ubiquitin riboprobe.

Library Screening, Subcloning, and Sequencing. *Arabidopsis* genomic libraries (9- to 15-kb fraction of *Sau*3A-digested C24 genomic DNA in λEMBL4) were screened with random primed probes (11). Screening and subsequent subcloning involved standard conditions and DNA manipulation, as described in Ausubel *et al.* (11). Automated DNA sequencing, with both dye primer and dye terminator reactions, was performed according to company specifications (Perkin–Elmer).

Reverse Transcriptase-PCR Analysis. Reverse transcriptase-PCR was carried out on first strand cDNA that was synthesized from RNA with Moloney murine leukemia virus reverse transcriptase as per company specifications (Promega). PCR was then performed with specific primers (100 nM) and standard PCR conditions (11). The resulting reverse transcriptase-PCR products were cloned and sequenced, confirming that the predicted introns are absent from the final transcript.

Southern Analysis. Genomic DNA was extracted using the method of Taylor and Powell (12). Hybridization and post hybridization treatment was as described in Ausubel *et al.* (11).

Database Search. Database searches were conducted through key word searching of the Expressed Sequence Tag database on the National Center for Biotechnology World Wide Web site (http://www.ncbi.nlm.nih.gov).

Expression, Purification, and Analysis Of Recombinant Proteins. cDNA clones of *AHB1* and *AHB2* were fused to an efficient ribosomal binding site, as per Springer and Sligar (13), and inserted into the pET24(+) expression vector (Novagen). *Escherichia coli*, strain BL21(DE3), containing each expression vector was grown to OD_{600nM} of 0.6, and then recombinant protein expression was induced with 1 mM isopropyl β-D-thiogalactoside. After 4 h at 37°C, cells were centrifuged and resuspended in 100 mM Tris·HCl (pH 8.0), sonicated, and cleared by centrifugation; 50 mg/ml hemin·HCl was added to the soluble fraction. Recombinant AHB1 protein was purified

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database [accession numbers U94998 (AHB1) and U94999 (AHB2)].

§Present address: Department of Biology, Texas A & M University, College Station, TX 77843-2133.
‖To whom reprint requests should be addressed. e-mail: liz@pican.pi.csiro.au.

a

```
                                                       I1 ──►                            70
                                                        *$      C2        CD1             E7
Barley .SAAEGAVVF SEEKEALVLK SWAIMKKDSA NLGLRFFLKI FEIAPSARQM FPFLRDSDVP LETNPKLKTH
Casnon .SSTLEGRGF TEEQEALVVK SWSAMKPNAG ELGLKFFLKI FEIAPSAQKL FSFLKDSNVP LERNPKLKSH
Soynon .TTTLE.RGF SEEQEALVVK SWNVMKKNSG ELGLKFFLKI FEIAPSAQKL FSFLRDSTVP LEQNPKLKPH
Parasp .SSSEVNKVF TEEQEALVVK AWAVMKKNSA ELGLQFFLKI FEIAPSAKNL FSYLKDSPVP LEQNPKLKPH
AHB1   ..ESEGKIVF TEEQEALVVK SWSVMKKNSA ELGLKLFIKI FEIAPTTKKM FSFLRDSPIP AEQNPKLKPH
AHB2   .....GEIGF TEKQEALVKE SWEILKQDIP KYSLHFFSQI LEIAPAAKGL FSFLRDSDEV PHNNPKLKAH
Cassym ........AL TEKQEALLKQ SWEVLKQNIP AHSLRLFALI IEAAPESKYV FSFLKDSNEI PENNPKLKAH
Soylba .......GAF TEKQEALVSS SFEAFKANIP QYSVVFYTSI LEKAPAAKDL FSFL--ANGV DPTNPKLTGH

              I2                                    I3        ◄──                140
              *                         F8          *
Barley AVSVFVMTCE AAAQLRKAGK ITVRETTLKR LGGTHLKYGV ADGHFEVTRF ALLETIKEAL PADMWGPEMR
casnon AMSVFLMTCE SAVQLRKAGK VTVRESSLKK LGASHFKHGV ADEHFEVTKF ALLETIKEAV P-ETWSPEMK
soynon AVSVFVMTCD SAVQLRKAGK VTVRESNLKK LGATHFRTGV ANEHFEVTKF ALLETIKEAV P-EMWSPAMK
Parasp ATTVFVMTCE SAVQLRKAGK VTVKESDLKR IGAIHFKTGV VNEHFEVTRF ALLETIKEAV P-EMWSPEMK
AHB1   AMSVFVMCCE SAVQLRKTGK VTVRETTLKR LGASHSKYGV VDEHFEVAKY ALLETIKEAV P-EMWSPEMK
AHB2   AVKVFKMTCE TAIQLREEGK VVVADTTLQY LGSIHLKSGV IDPHFEVVKE ALLRTLKEGL G-EKYNEEVE
cassym AAVIFKTICE SATELRQKGH AVWDNNTLKR LGSIHLKNKI TDPHFEVMKG ALLGTIKEAI K-ENWSDEMG
soylba AEKLFALVRD SAGQLKASGT VV-AD-A--A LGSVHAQKAV TDPEFVVVKE ALLKTIKAAV G-DKWSDELA

                                163

Barley NAWGEAYDQL VAAIKQEMKP AE.
casnon NAWGEAYDKL VAAIKLEMKP SS.
soynon NAWGEAYDQL VDAIKSEMKP PSS
Parasp NAWGVAYDQL VAAIKFEMKP SST
AHB1   VAWGQAYDHL VAAIKAEMNL SN.
AHB2   GAWSQAYDHL ALAIKTEMKQ EES
cassym QAWTEAYNQL VATIKAEMK. ...
soylba RAWEVAYDEL AAAIKK.... ...
```

b **SIMILARITY** **PARSIMONY**

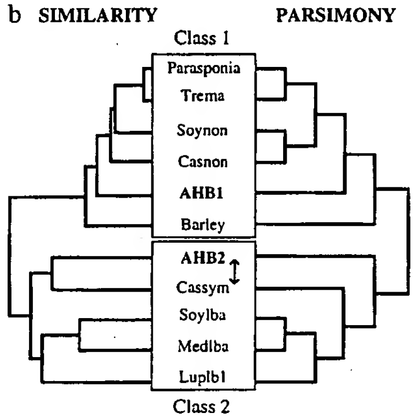

FIG. 1. (*a*) Alignment of some plant hemoglobin protein sequences. Proximal and distal histidines, phenylalanine CD1, and proline C2 are shown in bold. The location of primers used to amplify *AHB1* PCR fragment are shown as arrows. $ represents the 5′ end of the *AHB2* cDNA (Clone ID 11161). Intron positions are marked with stars and labeled I1-I3. Class 1 and 2 hemoglobins are separated by the horizontal line. (*b*) Cladistic analysis of plant hemoglobins. The left tree is a similarity tree of plant hemoglobin protein sequences created using the Genetics Computer Group PILEUP program. Class 1 and 2 hemoglobins form two distinct clades. The right tree is one of the two most parsimonious trees (CI .855, RC .654) generated by an exhaustive search of all possible trees using the PAUP, Ver. 3.1.1 program. The two trees of maximum parsimony differed in whether they grouped the *Casuarina* symbiotic hemoglobin or AHB2 with the leghemoglobins (represented by the arrow). The bootstrap value for the basal branch of this tree is 100%. The two classes of plant hemoglobins are highlighted by boxes. Hemoglobins (and GenBank accession numbers): barley: barley nonsymbiotic hemoglobin (U01228); casnon: *Casuarina* nonsymbiotic hemoglobin (X53950); soynon: soybean nonsymbiotic hemoglobin (U47143); parasp: *Parasponia* hemoglobin (U27194); AHB1: This study (U94998); AHB2: this study (U94999); cassym: *Casuarina* symbiotic hemoglobin (77695); soylba: soybean leghemoglobin A (J01299); medslba: *Medico sativa* leghemoglobin A (X14311); luplb1: lupin leghemoglobin 1 (Y00401).

by 25–70% ammonium sulfate fractionation, resuspended in 20 mM Tris·HCl (pH 8.0), and dialyzed twice into 4 liters of the same buffer. Dialyzed extracts were concentrated (YM-10 membrane, Amicon) and passed over a cation exchange column (DE52 Whatman) in 20 mM Tris·HCl (pH 8.0). Red fractions were concentrated, adjusted to pH 6.0 with dilute acetic acid, and passed over an anion exchange column (CM52, Whatman) with 20 mM sodium phosphate buffer (pH 8.0). The red fraction was concentrated to 10 mg/ml and stored at −80°C. Recombinant AHB2 protein was purified as was described for AHB1, with the following modification: 25–90% ammonium sulfate fractionation. Elution from DE52 resin was achieved with 250 mM Tris·HCl, following 5 column volumes of 100 mM Tris·HCl wash buffer. After elution from CM52 resin, red fractions were adjusted to 25% ammonium sulfate saturation (pH 8.0) and passed over a phenol-Sepharose column (equilibrated with 20 mM Tris·HCl/25% ammonium sulfate) then eluted with 20 mM Tris·HCl/12.5% ammonium sulfate and concentrated to 10 mg/ml for storage. Rate constants for O_2 and CO binding were measured using rapid mixing and laser photolysis methods (14).

RESULTS

A Nonsymbiotic Hemoglobin Gene from *Arabidopsis*. Using degenerate primers that amplified a nonsymbiotic hemoglobin gene fragment from soybean (7), we amplified a 459-bp fragment of a nonsymbiotic hemoglobin gene from *Arabidop-*

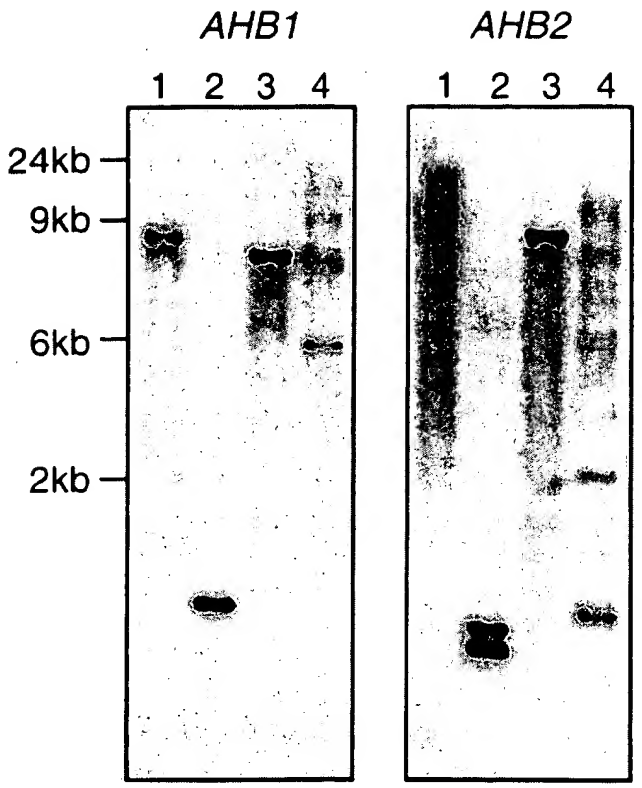

FIG. 2. Southern blots of *AHB1* and *AHB2*. *Arabidopsis* (C24 ecotype) genomic DNA digested with *Bam*HI, *Dra*I, *Eco*RI, and *Xba*I (lanes 1, 2, 3, and 4, respectively) and hybridized with *AHB1* and *AHB2* cDNA probes. (Extra bands on the *Xba*I lanes are due to partial digestion.)

sis. This gene fragment was used to isolate a full-length hemoglobin gene, *AHB1*, with a predicted amino acid sequence that shows strong homology to the sequences of nonsymbiotic hemoglobins from other plants (Fig. 1 *a* and *b*), including predicted ORFs containing amino acid residues that are conserved in all plant hemoglobins [the proximal F8 and distal E7 histidines, phenylalanine CD1 and proline C2 (15)]. There are three introns in positions identical to the introns occurring in all other plant hemoglobins (Fig. 1*a*). Southern analysis, at high stringency, suggests this gene is present in single or low copy number in the *Arabidopsis* genome (Fig. 2).

A Leghemoglobin-Like Gene from Arabidopsis. We identified a second hemoglobin gene (*AHB2*) by searching the *Arabidopsis* Expressed Sequence Tag database. The single partial length cDNA showing homology to plant hemoglobins (clone ID 11161) does not correspond in sequence to the *AHB1* gene, with only 69% identity at the nucleotide level. This gene, *AHB2*, has three introns in identical positions to those of other plant hemoglobins and, like *AHB1*, appears to be a single copy gene according to Southern blot data (*Dra*I cuts within the coding sequence of *AHB2*, producing two bands in this lane; Fig. 2). When the amino acid sequence of AHB2 is aligned with other plant hemoglobin protein sequences, it appears to be more similar to the symbiotic hemoglobins of *Casuarina* and legumes than it does to nonsymbiotic plant hemoglobins (Fig. 1*a*). If the AHB2 protein sequence is used in a BLAST (basic local alignment search tool) search, it consistently detects symbiotic hemoglobins with higher BLAST scores than most nonsymbiotic hemoglobins. A similarity tree of plant hemoglobin protein sequences, constructed with the PILEUP function of the Genetics Computer Group package, also shows AHB2 to be more similar to the symbiotic hemoglobins of legumes and *Casuarina* (Fig. 1*b*). The same clade occurs in the two trees of maximum parsimony constructed using the PAUP program, Ver. 3.1.1, one of which is presented (Fig. 1*b*). The bootstrap value for the basal branch of this clade is high (100%), supporting the grouping of AHB2 with the symbiotic hemoglobins of legumes and *Casuarina*.

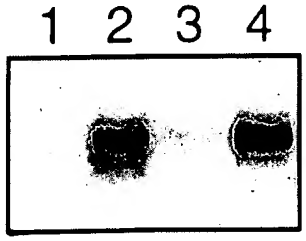

FIG. 3. AHB1 expression in root and rosette tissues. Total RNA extracted from rosette leaves (lane 1), rosettes treated with 5% oxygen (lane 2), roots (lane 3), and roots treated with 5% oxygen (lane 4) hybridized with an *AHB1* antisense riboprobe.

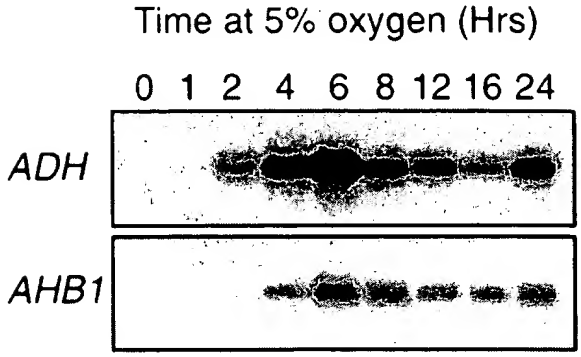

FIG. 4. Time course analysis of the *AHB1* hypoxia response. Total RNA from whole plants treated with 5% oxygen for varying times hybridized with either *AHB1* or *ADH* antisense riboprobes.

AHB1 Is Expressed in Root Tissue and Is Hypoxia-Inducible. *AHB1* expression, examined with RNA gel blots, is detectable at low levels in root tissue and is strongly induced in both the roots and rosette leaves of plants subjected to hypoxic conditions (Fig. 3). The time course of the *AHB1* response to low oxygen levels is similar to that of the alcohol dehydrogenase gene (*ADH*), a well characterized hypoxic response gene (ref. 16; Fig. 4). The maximum level of *AHB1* transcript was observed after 6–8 h of hypoxia treatment. Transcription of *AHB1* is induced under a 5% oxygen environment, but stronger induction occurs at 0.1% oxygen (Fig. 5). A transcriptional response also was elicited by a 1% sucrose solution. Cold, dehydration, heat shock, oxidative stress, and wounding treatments did not affect *AHB1* transcript levels.

AHB2 Is Expressed in Rosette Leaves. A low level of *AHB2* expression was found in rosette leaves, but unlike *AHB1*, there

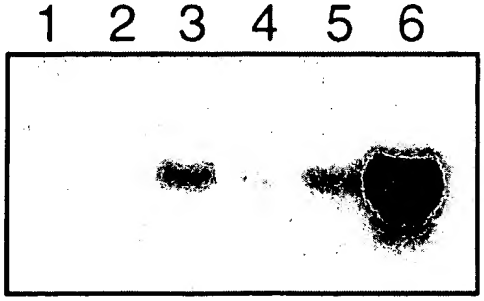

FIG. 5. AHB1 expression during environmental stress treatments. Total RNA from control plants (lane 1) or plants subjected to chilling (lane 2), 1% sucrose (lane 3), dehydration (lane 4), 5% oxygen (lane 5), or 0.1% oxygen (lane 6) hybridized with an *AHB1* antisense riboprobe.

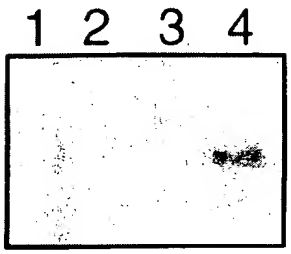

FIG. 6. Induction of *AHB2* by low temperature. Total RNA from root tissue (lane1) or rosette leaves (lane 2) and root tissue after chilling treatment (lane 3) or rosette leaves after chilling treatment (lane 4) hybridized with an *AHB2* antisense riboprobe.

was no detectable expression in root tissue. *AHB2* expression was not affected by sucrose or hypoxia treatment, but low temperature did increase *AHB2* transcript levels (Fig. 6). The low temperature response seems to be confined to the rosette leaves with no response evident in root tissue (Fig. 6). Wounding, heat shock, oxidative stress, and dehydration had no effect on the level of *AHB2* transcript.

Recombinant AHB1 and AHB2 Proteins Have Different Oxygen Binding Properties. Recombinant AHB1 protein, expressed in *E. coli*, was purified, and the kinetics of O_2 and CO binding were measured. AHB1 shows a remarkably high O_2 affinity, with a P_{50} value equal to 1.6 nM at pH 7, 20°C (Table 1). Similar oxygen binding parameters have been observed for the nonsymbiotic hemoglobins of soybean and rice (R.A.W., unpublished data, and M. Hargrove, personal communication). Lowering the pH to 5 raises the P_{50} to 6.7 nM (Table 1), but this value is still smaller than that of any known plant cytochrome oxidase (17), making it unlikely that AHB1 could transport oxygen to these terminal oxidases. AHB1 also has a low O_2 dissociation rate constant compared with that of soybean leghemoglobin A, a known oxygen transporter (Table 1). Thus, AHB1 would appear to be an oxygen-scavenging protein rather than an oxygen transporter.

In contrast, recombinant AHB2 has a much lower affinity for O_2, with P_{50} values of 130 nM at pH 7 and 2,700 nM at pH 5 (Table 1). The dissociation rate constant is low at pH 7 but increases 20-fold at pH 5, approaching the value for soybean leghemoglobin A under the same acidic conditions. The low association rate constant for O_2 binding to AHB2 appears to be caused by the reversible formation of a hexacoordinate complex (hemochrome) in deoxygenated AHB2. The imidazole side chain of the distal histidine (E7), the only polar residue in the distal pocket, may coordinate directly to the iron atom of the heme group, thereby inhibiting rapid O_2 binding and facilitating O_2 dissociation.

Table 1. Comparison of parameters* for O_2 binding to recombinant hemoglobins at 20°C

Protein	pH	P_{50}, nM	k_{O_2}, s^{-1}	M K_{CO}/K_{O_2}
AHB1	7.0	1.6	0.12	2.0
	5.0	6.7	0.38	
AHB2	7.0	130	0.14	26
	5.0	2700	2.1	
Soy lbA†	7.0	44	5.6	87
	5.0	20	2.7	

*Where k_{O_2} and k'_{O_2} are the bimolecular dissociation and association rate constants for O_2 binding, respectively. The dissolved O_2 concentration at which hemoglobin is half saturated (P_{50}) is equivalent to the dissociation equilibrium constant for O_2 ($K'_{O_2} = k_{O_2}/k'_{O_2}$). K_{O_2} and K_{CO} are the association equilibrium constants for O_2 and CO, respectively. M is the partition coefficient expressing the relative affinities for O_2 and CO.

†Recombinant soybean leghemoglobin A.

DISCUSSION

AHB1 is related to the nonsymbiotic hemoglobin genes present in other plants. The identification of this gene in *Arabidopsis* supports the idea that nonsymbiotic hemoglobin genes may be present in all plants (6). The expression pattern of *AHB1* is similar to that of other nonsymbiotic hemoglobins; low levels of expression in root tissue are common to all nonsymbiotic hemoglobins (3–8), sucrose inducibility has been observed for the *Casuarina* nonsymbiotic hemoglobin (4), and the hypoxic induction of *AHB1* resembles that of the barley hemoglobin (8).

AHB2 represents a new class of nonsymbiotic hemoglobin gene. It shows limited homology to the other nonsymbiotic hemoglobins and has a different pattern of gene expression. *AHB2*-like genes may be present in a wide range of plants. In support of this theory, we have cloned an *AHB2* homologue from *Brassica napus* (R.A.W., unpublished data). We suggest that the *AHB1*-like "nonsymbiotic hemoglobin" gene family be referred to as class 1 plant hemoglobin genes and that *AHB2*-like genes be referred to as class 2 plant hemoglobin genes.

It has been suggested that symbiotic hemoglobin genes of legumes and *Casuarina* arose from gene duplication of a nonsymbiotic (class 1) hemoglobin gene in the ancestor of nitrogen-fixing plants (5, 6). The similarity between *AHB2* and symbiotic hemoglobins suggests an alternative explanation; a class 2 hemoglobin gene may have been the direct ancestor of the symbiotic hemoglobin genes. In contrast, it appears that a class 1 hemoglobin has been recruited into symbiotic function in *Parasponia* nodules. The recruitment of either class 1 or class 2 plant hemoglobins into symbiotic roles seems to have been associated with changes in gene expression, leading to high levels of expression in nodules, and with alterations in the oxygen binding properties of the protein.

Leghemoglobins facilitate oxygen flux down an oxygen gradient without the involvement of other proteins. It is unlikely that either AHB1 or AHB2 acts as an oxygen transporter in this manner. The oxygen affinity of AHB1 is too high for release of oxygen to cytochrome oxidase, and the oxygen dissociation rates of both proteins are too slow for efficient oxygen transport. Instead, these proteins may act as oxygen scavengers, binding oxygen until interactions with other proteins occur. The eventual destination for bound oxygen could be mitochondrial metabolism or oxygen-consuming enzymatic reactions. It is also possible that AHB1 or AHB2 have unidentified catalytic properties. Analysis of transgenic plants that over- or underexpress AHB1 or AHB2, combined with cellular localization of these proteins, should lead to a better understanding of the functions of these and other plant hemoglobins.

We thank Janice Norman for her excellent technical assistance and Cyril Appleby for his helpful discussions.

1. Kubo, H. (1939) in Appleby, C. A. (1992) *Sci. Prog.* **76,** 365–398.
2. Appleby, C. A. (1984) *Annu. Rev. Plant Physiol.* **35,** 443–478.
3. Fleming, A. I., Wittenberg, J. B., Wittenberg, B. A., Dudman, W. F. & Appleby, C. A. (1987) *Biochim. Biophys. Acta* **911,** 209–220.
4. Jacobsen-Lyon, K., Jensen, E. Ø., Jorgensen, J., Marcker, K. A., Peacock, W. J. & Dennis, E. S. (1995) *Plant Cell* **7,** 213–223.
5. Landsmann, J., Dennis, E. S., Higgins, T. J., Appleby, C. A., Kortt, A. A. & Peacock, W. J. (1986) *Nature (London)* **324,** 166–168.
6. Bogusz, D., Appleby, C. A., Landsmann, J., Dennis, E. S. & Peacock, W. J. (1988) *Nature (London)* **331,** 178–180.
7. Andersson, C. R., Jensen, E. Ø., Llewellyn, D. J., Dennis, E. S. & Peacock, W. J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 5682–5687.
8. E. R. Taylor, E. R., Nie, X. Z., MacGregor, A. W. & Hill, R. D. (1994) *Plant Mol. Biol.* **24,** 853–862.
9. Dolferus, R., Jacobs, M., Peacock, W. J. & Dennis, E. S. (1994) *Plant Physiol.* **105,** 1075–1087.

10. Logemann, J., Schell, J. & Willmitzer, L. (1987) *Anal. Biochem.* **163,** 16–20.
11. Ausubel, F. M., Brenton, R., Kingston, R. E., Moore, D. D., Siedman, J. G., Smyth, J. A. & Struhl, K. (1994) *Current Protocols in Molecular Biology* (Wiley, New York).
12. Taylor, A. & Powell, B. (1982) *BRL Focus* **4,** 4.
13. Springer, B. A. & Sligar, S. G. (1997) *Proc. Natl. Acad. Sci. USA* **84,** 8961–8965.
14. Hargrove, M. S., Barry, J. K., Bruckner, E. A., Berry, M. B., Phillips G. N., Jr., Olson. J. S., Arrendo-Peter, R., Dean, J. M., Klucas & R. V. Sarath. G. (1997) *J. Mol. Biol* **266,** 1032–1042.
15. Chang, C. & Meyerowitz, E. M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 1408–1412.
16. Bashford, D., Chothia, C. & Lesk A. M. (1987) *J. Mol. Biol* **196,** 199–216.
17. Millar, A. H., Bergerson, F. J. & Day, D. A. (1994) *Plant Physiol. Biochem.* **32,** 847–852.
18. Wittenberg, J. B. & Wittenberg, B. A (1990) *Annu. Rev. Biophys. Biophys. Chem.* **19,** 217–241.

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 5682–5687, June 1996
Plant Biology

A new hemoglobin gene from soybean: A role for hemoglobin in all plants

(nonsymbiotic/leghemoglobin/evolution)

CAROL R. ANDERSSON*†‡, ERIK OSTERGAARD JENSEN§, DANNY J. LLEWELLYN*, ELIZABETH S. DENNIS*, AND W. JAMES PEACOCK*¶

*Commonwealth Scientific and Industrial Research Organization Plant Industry, G.P.O. Box 1600, Canberra, ACT 2601, Australia; †Division of Biochemistry and Molecular Biology, Australian National University, G.P.O. Box 4, Canberra ACT 2601, Australia; and §Laboratory of Gene Expression, University of Aarhus, Gustav Weids Vej 10, Aarhus C 8000 Denmark

Contributed by W. James Peacock, January 24, 1996

ABSTRACT **We have isolated a new hemoglobin gene from soybean. It is expressed in cotyledons, stems of seedlings, roots, young leaves, and in some cells in the nodules that are associated with the nitrogen-fixing *Bradyrhizobium* symbiont. This contrasts with the expression of the leghemoglobins, which are active only in the infected cells of the nodules. The deduced protein sequence of the new gene shows only 58% similarity to one of the soybean leghemoglobins, but 85–87% similarity to hemoglobins from the nonlegumes *Parasponia*, *Casuarina*, and barley. The pattern of expression and the gene sequence indicate that this new gene is a nonsymbiotic legume hemoglobin. The finding of this gene in legumes and similar genes in other species strengthens our previous suggestion that genomes of all plants contain hemoglobin genes. The specialized leghemoglobin gene family may have arisen from a preexisting nonsymbiotic hemoglobin by gene duplication.**

Despite the common perception of hemoglobin as a blood protein, the protein is also found in many invertebrates, bacteria, fungi, and in some higher plants. Plant hemoglobin was thought to be restricted to the nitrogen-fixing root nodules of legumes in the highly specialized symbiosis with strains of *Rhizobium* or *Bradyrhizobium* bacteria. The function of these legume hemoglobins (leghemoglobins) is the facilitated diffusion of oxygen to the respiring bacteroids within the root nodule (for review, see ref. 1).

More recently, hemoglobin and hemoglobin genes have been discovered in the nonlegume genera *Parasponia*, *Trema*, *Casuarina*, *Hordeum*, *Triticum*, and *Zea* (2–8). The finding of hemoglobin genes in non-nodulating plants (*Trema*, *Hordeum*, *Triticum*, and *Zea*) and in phylogenetically diverse plant genera, including monocots, reinforces the proposal that hemoglobin may be present in all plants (9). Two roles for plant hemoglobin outside the nitrogen-fixing symbiosis have been proposed: (*i*) as a facilitator of oxygen diffusion in rapidly respiring cells or (*ii*) as an oxygen sensor involved in switching plant metabolism to anaerobic pathways (9).

In general terms, there appear to be two different types of plant hemoglobins, the symbiotic (nodule specific and associated with a symbiotic interaction involving microorganisms) and the nonsymbiotic hemoglobins (expressed in non-nodule tissues such as roots and stems). The situation in the nodulating tree *Parasponia andersonii* is unusual in that a single gene encodes a protein that has both a symbiotic and nonsymbiotic role (3).

The symbiotic hemoglobin genes have been postulated to have arisen by duplication of preexisting, nonsymbiotic, hemoglobin genes (9). *Casuarina glauca* presents a possible example of such a gene duplication and divergence process in that it has both a single nonsymbiotic hemoglobin gene expressed in all tissues analyzed, and a different, symbiotic hemoglobin gene family expressed exclusively in the nodules induced by *Frankia*, an actinomycete (6, 7). Gene duplication with subsequent specialization is a common evolutionary strategy and recent evidence suggests that in legumes other nodule specific proteins have arisen by this mechanism, e.g., nodulin 26 (10).

If symbiotic hemoglobin genes were derived by duplication from nonsymbiotic genes, then where are the nonsymbiotic hemoglobins in legume species? Legumes have well-characterized, nodule-specific leghemoglobin gene families, but apart from a single report of a hemoglobin-like protein in the seeds of winged bean (11), expression of hemoglobin genes outside nodules has never been detected by either nucleic acid hybridization or reporter gene assays.

In this paper, we report the isolation and characterization of a nonsymbiotic hemoglobin gene from soybean and confirm that similar genes are present in other legumes. The gene is clearly related to nonsymbiotic plant hemoglobins on the basis of its predicted amino acid sequence and gene structure, including the promoter sequence; its expression pattern within the plant is unlike that of the symbiotic leghemoglobins.

MATERIALS AND METHODS

Isolation of Genomic DNA and Southern Blot Analysis. Isolation of genomic DNA and Southern blot analysis of hemoglobin genes from *Glycine max* cv Lincoln, *Pisum sativum* cv Greenfeast, *Pisum humile*, *Medicago sativa* cv R15 and *Trifolium repens* cv Haifa was as described (3). Probes were either the PCR fragment soyhb1f+4r2 or the leghemoglobin clones 13.0/1.6 and 7.5/2.0 (12), which together cover the entire coding region of soybean *lbps1*.

PCR Reactions. Degenerate primers were designed to conserved regions of plant nonsymbiotic hemoglobins (see Fig. 1) as follows: Hbexon1f (hemoglobin exon 1 forward primer), 5′-CGGAATTCGA(A/G)GA(A/G)(C/G)A(A/G)GA(A/G)GCI(T/C)TIGT; Hbexon2f (exon2 forward primer), 5′-CGGAATTCATITT(T/C)GA(A/G)ATIGCICC; Hbexon2r (exon 2 reverse primer), 5′-CGGGATCCGC(A/G)TGIII(T/C)TTIA(A/G)(T/C)TTIGG(A/G)TT; Hbexon4r2 (exon 4 reverse primer), 5′- CGGGATCCGC(T/C)TC(T/C)TTIATIGT(T/C)TC.

Template DNA was digested with *Hin*dIII or *Xho*I, followed by heat inactivation at 65°C for 20 min. The PCR reaction included *Taq* DNA polymerase buffer (Promega)/1 mM

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Data deposition: The sequence reported in this paper has been deposited in the GenBank data base (accession no. U47143).

‡Present address: Department of Biology, Texas A&M University, College Station, TX 77843.
¶To whom reprint requests should be addressed. e-mail: jpeacock@pi.csiro.au.

$MgCl_2$/200 μM each dNTP/1 μM each primer/250 ng digested genomic DNA/2.5 units *Taq* DNA polymerase (Promega) in a total reaction volume of 15 μl. Reactions were performed using a Corbett thermal cycler. The reaction incorporated stepwise reductions in annealing temperature ("touchdown PCR") from 50°C to 37°C, and slow ramp times from the annealing to the extension temperature. The denaturation times were 30 sec, with annealing and extension times of 1 min. PCR products were digested with *Bam*HI and *Eco*RI, purified from a polyacrylamide gel, and cloned into pBluescript SK (−) (Stratagene) or pGem 3Zf(−) (Promega). Plasmid inserts were sequenced using a Pharmacia T7 sequencing kit.

The end-labeled hemoglobin specific primer soyhb2f35 (soybean hemoglobin exon 2 forward 35 mer) 5'- CTCATTCTTGAGAGATTCAACGGTTCCTTTGGAGC was used to identify longer products in subsequent PCR reactions.

Isolation of Soybean Genomic Clones. A soybean genomic library in lambda EMBL4 was constructed from size fractionated, partial *Sau*3A digested genomic DNA and screened with a soyhb1f+4r2 probe using standard methods (13).

Northern Blot Analysis. Soybean tissue from shoots, leaves, stems, flowers (all stages), seed pods, roots, nodules, and germinating seeds were harvested, snap frozen in liquid nitrogen, and stored at −80°C. To eliminate the possibility that root tissue contained nodules, roots were harvested from seedlings grown from surface sterilized seeds sown in sterilized vermiculite/sand. Total RNA was isolated and analyzed by Northern blotting as described (3). Filters were probed with an antisense RNA probe (Amersham *in vitro* transcription kit) synthesized from soyhb1f+4r2, subsequently stripped and reprobed with an *A. thaliana* ubiquitin (14) antisense RNA probe. The hybridization signals were quantified using a PhosphorImager (Fujix) and the abundance of the soybean hemoglobin message expressed relative to ubiquitin. The size of transcripts were estimated relative to RNA standards (BRL).

```
                1                  hbexon1f>                    *hbexon2f>     60
      Tremahb  .......SSS EVDKVFTEEQ EALVVKSWAV MKKNSAELGL KFFLKIFEIA PSAKNLFSYL
       Parahb  .......SSS EVNKVFTEEQ EALVVKAWAV MKKNSAELGL QFFLKIFEIA PSAKNLFSYL
   Casnonsymh  ........ST LEGRGFTEEQ EALVVKSWSA MKPNAGELGL KFFLKIFEIA PSAQKLFSFL
        Soyhb  ........TT TLERGFSEEQ EALVVKSWNV MKKNSGELGL KFFLKIFEIA PSAQKLFSFL
      Ricehb1  ALVEDNNAV. ..AVSFSEEQ EALVLKSWAI LKKDSANIAL RFFLKIFEVA PSASQMFSFL
      Ricehb2  ALVEGNNGVS GGAVSFSEEQ EALVLKSWAI MKKDSANIGL RFFLKIFEVA PSGCQMFSFL
      Barleyhb .......SAA EGAVVFSEEK EALVLKSWAI MKKDSANLGL RFFLKIFEIA PSARQMFPFL
      Soylbc2  .......... ...GAFTEKQ EALVSSSFEA FKANIPQYSV VFYTSILEKA PAAKDLFSFL
      Soylbc3  .......... ...GAFTDKQ EALVSSSFEA FKTNIPQYSV VFYTSILEKA PVAKDLFSFL
       Soylba  .......... ...VAFTEKQ DALVSSSFEA FKANIPQYSV VFYTSILEKA PAAKDLFSFL
      Soylbc1  .......... ...GAFTEKQ EALVSSSFEA FKANIPQYSV VFYNSILEKA PAAKDLFSFL
    Kidneybean .......... ...GAFTEKQ EALVNSSWEA FKGNIPQYSV VFYTSILEKA PAAKNLFSFL
    Sesbaniaii .......... ....GFTDKQ EALVNASYEA FKKNLPGHSV LFYSFILEKE PAAKGLFSFL
     Medicago1 .......... ....SFTDKQ EALVNSSYEA FKQNLSGYSV FFYTVILEKA PAAKGLFSFL
     Broadbean .......... ....GFTEQQ EALVNSSSQL FKQNPSNYSV LFYTIILQKA PTAKAMFSFL
        Pealb  .......... ....GFTDKQ EALVNSSSE. FKQNLQGYAT LFYTIILEKA PAAKGLFSFL
     Lupin11b  .......... ...GVLTDVQ VALVKSSFEE FNANIPKNTH RFFTLVLEIA PGAKDLFSFL
     Cassymhb  .......... ....ALTEKQ EALLKQSWEV LKQNIPAHSL RLFALILEAA PESKYVFSFL

                61      <hbexon2r      *                                    120
      Tremahb  KDSPRPLEQN PKLKPHAMTV FVMTCESAVQ LRKAGKVTVR ESNLKRLGAI HFKNGVVNEH
       Parahb  KDSPVPLEQN PKLKPHATTV FVMTCESAVQ LRKAGKVTVK ESDLKRIGAI HFKTGVVNEH
   Casnonsymh  KDSNVPLERN PKLKSHAMSV FLMTCESAVQ LRKAGKVTVR ESSLKKLGAS HFKHGVADEH
        Soyhb  RDSTVPLEQN PKLKPHAVSV FVMTCDSAVQ LRKAGKVTVR ESNLKKLGAT HFRTGVANEH
      Ricehb1  RNSDVPLEKN PKLKTHAMSV FVMTCEAAAQ CGKAGKVTVR DTTLKRLGAT XLKYGVGXPX
      Ricehb2  RNSDVPLEKN PKLKTHGMSV FVMTCEAAAX CGKAGKVTVR DTTLKRLGAT ..........
      Barleyhb RDSDVPLETN PKLKTHAVSV FVMTCEAAAQ LRKAGKITVR ETTLKRLGGT HLKYGVADGH
      Soylbc2  ..SNGVDPSN PKLTGHAEKL FGLVRDSAGQ LKANGTVV.A .DA..ALGSI HAQKAITDPQ
      Soylbc3  ..ANGVDPTN PKLTGHAEKL FGLVRDSAGQ LKASGTVV.I .DA..ALGSI HAQKAITDPQ
       Soylba  ..ANGVDPTN PKLTGHAEKL FALVRDSAGQ LKASGTVV.A .DA..ALGSV HAQKAVTDPQ
      Soylbc1  ..ANGVDPTN PKLTGHAEKL FALVRDSAGQ LKTNGTVV.A .DA..ALVSI HAQKAVTDPQ
    Kidneybean ..ANGVDPTN PKLTAHAESL FGLVRDSAAQ LRANGAVV.A .DA..ALGSI HSQKGVNDSQ
    Sesbaniaii KDSDGVPQNN PSLQAHAEKV FGLVHDAAGQ LRATGVVVLA .DA..SLGSV HVQKGVTDPH
     Medicago1 KDSAGV.QDS PQLQAHAEKV FGLVRDSASQ LRATGGVVLG .DA..ALGAI HIQKGVVDPH
     Broadbean KDSAGVV.DS PKLGAHAEKV FGMVRDSAVQ LRATGEVVL. .DG..KDGSI HIQKGVLDPH
        Pealb  KDTAGV.EDS PKLQAHAEQV FGLVRDSAAQ LRTKGEVVLG .NA..TLGAI HAGKGVTNPH
     Lupin11b  KGSSEVPQNN PDLQAHAGKV FKLTYEAAIQ LQVNGAVASD .ATLKSLGSV HVSKGVVDAH
     Cassymhb  KDSNEIPENN PKLKAHAAVI FKTICESATE LRQKGHAVWD NNTLKRLGSI HLKNKITDPH

                121     <hbexon4r2                          169
      Tremahb  FEVTRFALLE TIKEAVP.EM WSPEMKNAWG EAYDQLVAAI KSEMKPSST
       Parahb  FEVTRFALLE TIKEAVP.EM WSPEMKNAWG VAYDQLVAAI KFEMKPSST
   Casnonsymh  FEVTKFALLE TIKEAVP.ET WSPEMKNAWG EAYDKLVAAI KLEMKPSS.
        Soyhb  FEVTKFALLE TIKEAVP.EM WSPAMKNAWG EAYDQLVDAI KSEMKPPSS
      Ricehb1  FEVVKFALLD TIKEEVPADM WSPAMKSAWS EAYDHLVAAI KQEMKPAE.
      Ricehb2  .......... .......... .......... .......... .........
      Barleyhb FEVTRFALLE TIKEALPADM WGPEMRNAWG EAYDQLVAAI KQEMKPAE.
      Soylbc2  FVVVKEALLK TIKEAVG.DK WSDELSSAWE VAYDELAAAI KKAF.....
      Soylbc3  FVVVKEALLK TIKEAVG.DK WSDELSSAWE VAYDELAAAI KKAF.....
       Soylba  FVVVKEALLK TIKAAVG.DK WSDELSRAWE VAYDELAAAI KKA......
      Soylbc1  FVVVKEALLK TIKEAVG.GN WSDELSSAWE VAYDELAAAI KKA......
    Kidneybean FLVVKEALLK TLKEAVG.DK WTDELSTALE LAYDELAAAI KKAYA....
    Sesbaniaii FVVVKEALLK TLKEAAG.AT WSDEVSEAWE VAYDGLAAAI KKAMS....
     Medicago1 FVVVKEALLK TIKEAAG.DK WSEELSTAWE VAYDALATEI KKAMS....
     Broadbean FVVVKEALLK TIKEASG.DK WSEELSAAWE VAYDGLATAI KAA......
        Pealb  FVVVKEALLQ TIKKASG.NN WSEELNTAWE VAYDGLATAI KKAMKTA..
     Lupin11b  FPVVKEAILK TIKEVVG.DK WSEELNTAWT IAYDELAIII KKEMKDAA.
     Cassymhb  FEVMKGALLG TIKEAIK.EN WSDEMGCAWT EAYNQLVATI KAEMKE...
                *
```

FIG. 1. Alignment of the predicted amino acid sequences of various plant hemoglobins using the GCG PILEUP program. The highly conserved residues involved in heme and ligand binding [the distal and proximal histidines, phenyalanine CD1 and proline C2 (1)] are in bold. Intron positions are marked with asterisks. Conserved regions used to design degenerate PCR primers are overlined. Ricehb1 and ricehb2, expressed sequence tag sequences from rice, GenBank accession numbers D15507and D23324 or D25122 and D22678, respectively (X is indicated at positions of ambiguity in the entered DNA sequence); barleyhb (8); *Trema tomentosa* hemoglobin (tremahb) (4); *Parasponia andersonii* hemoglobin (parahb) (3), *Casuarina glauca* nonsymbiotic hemoglobin (casnonsymhb) (6); soybean nonsymbiotic hemoglobin (soyhb) (this study), lupin hemoglobin 1 (lupin1lb) (15); soybean lba and lbc1 leghemoglobins (soylba and soylbc1) (16); soybean lbc2 and lbc3 leghemoglobins (soylbc2 and soylbc3) (17); kidney bean leghemoglobin (kidneybean) (18); *Sesbania rostrata* leghemoglobin II (sesbaniaii) (19); *Medicago truncatula* leghemoglobin1 (medicago1) (20); broadbean leghemoglobin (broadbean) (21); pea leghemoglobin (pealb) (22), *Casuarina glauca* symbiotic hemoglobin (cassymhb) (7).

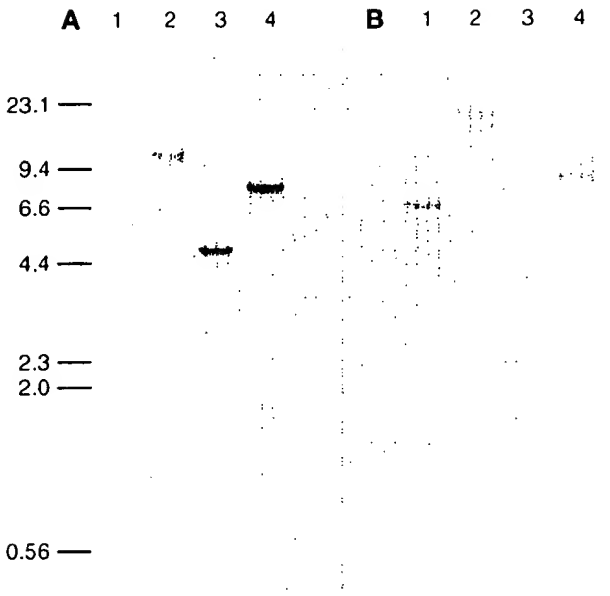

FIG. 2. Southern blot analysis of the soybean nonsymbiotic hemoglobin gene and leghemoglobin genes. *Glycine max* cv Lincoln genomic DNA was digested with *Eco*RI (lane 1), *Bam*HI (lane 2), *Hin*dIII, (lane 3) and *Bgl*II (lane 4). (*A*) Soybean nonsymbiotic hemoglobin (soybean 1f+4r2 PCR fragment probe). (*B*) Soybean leghemoglobins (*lbps1* 13.0/1.6 and 7.5/2.0 probes). Lambda *Hin*dIII size standards are in kb.

RESULTS

A New Hemoglobin Gene in the Soybean Genome. Short degenerate oligonucleotides designed from the conserved regions of known nonsymbiotic plant hemoglobins (Fig. 1) were used in a PCR reaction on soybean genomic DNA. Primers within the second exon (hbexon2f and hbexon2r) amplified a number of fragments including one of ≈110 bp, the size predicted from other plant hemoglobins. The 110-bp fragment had an open reading frame that contained residues known to be critical in hemoglobin structure and function, and was partially homologous to the known nonsymbiotic plant hemoglobins. A longer gene fragment was generated using degenerate primers derived from sequences in the first and last exons of nonsymbiotic hemoglobins (hbexon1f and hbexon4r2). As the genomic DNA was likely to include three introns, the hemoglobin-specific Soyhb2f35 probe was used to identify the correct product from this second PCR reaction.

The longer clone (1f+4r2) contained the original 110-bp fragment and was used to isolate corresponding genomic segments. The gene sequence (GenBank accession no. U47143) had a five amino acid N-terminal extension similar to that in the nonsymbiotic and *P. andersonii* hemoglobins (Fig. 1), and three introns at positions found in all plant hemoglobins.

The New Hemoglobin Gene Does not Hybridize to the Leghemoglobins; Homologues Are Present in Other Legumes. Southern blot analysis, using the 1f+4r2 PCR fragment and the previously described leghemoglobin gene probes, showed no hybridizing bands in common between these two classes of genes in soybean (Fig. 2) as expected from the low similarity between their nucleotide sequences (Table 1). Using the 1f+4r2 probe, *Eco*RI-digested soybean DNA displayed two hybridizing bands of about equal intensity, whereas the *Bam*HI and *Bgl*II lanes each had a single, more intense band. The *Hin*dIII digest produced two bands of different intensities. As the soybean 1f+4r2 probe does not contain an *Eco*RI site, we conclude that the soybean genome contains two copies of the new hemoglobin gene and, because there is a single *Bam*HI fragment, that these might be physically linked. Alternatively, and perhaps more likely, because soybean is an ancient tetraploid, the two copies may be present on homeologous chromosomes, giving rise to separate restriction fragments of equivalent length. It is unlikely that the two bands in the *Eco*RI and *Hin*dIII digests represent restriction fragment length polymorphisms of a single gene because of the inbred nature of the soybean cultivar.

Southern blot analysis of other legumes using the soybean 1f+4r2 probe at moderate hybridization stringency detected one or more cross-hybridizing bands in white clover, lucerne, and pea (data not shown), indicating that an homologous nonsymbiotic hemoglobin gene(s) is also present in the genomes of these species.

The Promoter of the New Hemoglobin Gene Does Not Contain the Nodulin Consensus Sequences. The symbiotic hemoglobin promoters contain two motifs that have been shown to be critical for nodule specific expression (the nodulin boxes, CTCTT and AAGAT) separated by six or seven nucleotides (23, 24). The nonsymbiotic (or dual function hemoglobin, in the case of the *P. andersonii* gene) hemoglobin genes from *C. glauca*, *Trema tomentosa*, and *P. andersonii* do not have these conserved motifs at this spacing (see Fig. 5). The critical CTCTT motif is absent from the nonsymbiotic genes. The promoter of the new soybean hemoglobin gene does not contain this motif either, but has the sequence CTCCC, identical in sequence and position to a motif in *P. andersonii* (see Fig. 5). Another motif shown to be critical for expression of the *P. andersonii* promoter (unpublished data), GAAGAG, is present in a similar position as GAAGGG is in the soybean promoter. These data suggest that the new soybean hemoglobin has a promoter that resembles promoters of the nonsymbiotic hemoglobins of other plant genera, more than it resembles the promoters of the symbiotic leghemoglobins in its own genome.

The New Soybean Hemoglobin Gene Is Expressed in Many Tissues. Northern blot analysis of total RNA isolated from various soybean tissues detected a nonsymbiotic hemoglobin mRNA of 830 nt. The message was readily detected in stems of mature plants and cotyledons of seedlings (Fig. 3). The gene was also expressed in roots, nodules, and young leaves of mature plants, but was virtually absent from older leaves, floral tissue, and seed pods. The level of expression in nodules was comparable to that in non-nodulated roots and considerably lower than that of leghemoglobin genes in the infected cells of nodules. The new hemoglobin's transcript was present at a

Table 1. Overall sequence similarity (%) between various plant hemoglobins

	Lupin1lb	Soylba	Cassymhb	Casnonsymhb	Parahb	Barleyhb	Soyhb
Lupin1lb		69	70	56	61	61	57
Soylba	72		67	61	62	61	58
Cassymhb	62	60		64	64	64	64
Casnonsymhb	59	54	56		84	84	87
Parahb	58	55	57	76		83	87
Barleyhb	56	60	56	69	69		85
Soyhb	58	53	56	76	78	69	

Numbers above the diagonal represent similarity between predicted hemoglobin amino acid sequences; numbers below the diagonal represent similarity between nucleotide sequences in the coding region (see Fig. 1 legend).

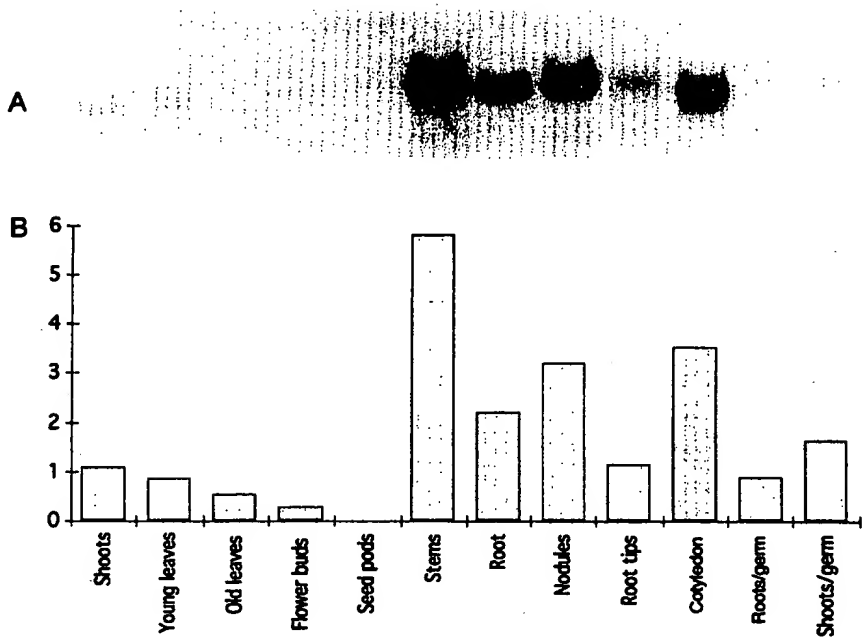

FIG. 3. Nonsymbiotic hemoglobin gene expression patterns in soybean seedlings and plants. Total RNA was isolated from soybean shoot tips, young leaves, old leaves, flower buds, whole seed pods, stems, roots, nodules, root tips, and cotyledons, roots, and shoots from germinating seeds. (*A*) Hybridization pattern with a riboprobe corresponding to the soybean PCR fragment 1f+4r2. (*B*) Rehybridization with an *Arabidopsis* ubiquitin probe provided levels of nonsymbiotic hemoglobin gene transcript relative to ubiquitin abundance.

lower concentration in root tips than in the remainder of the root (Fig. 3).

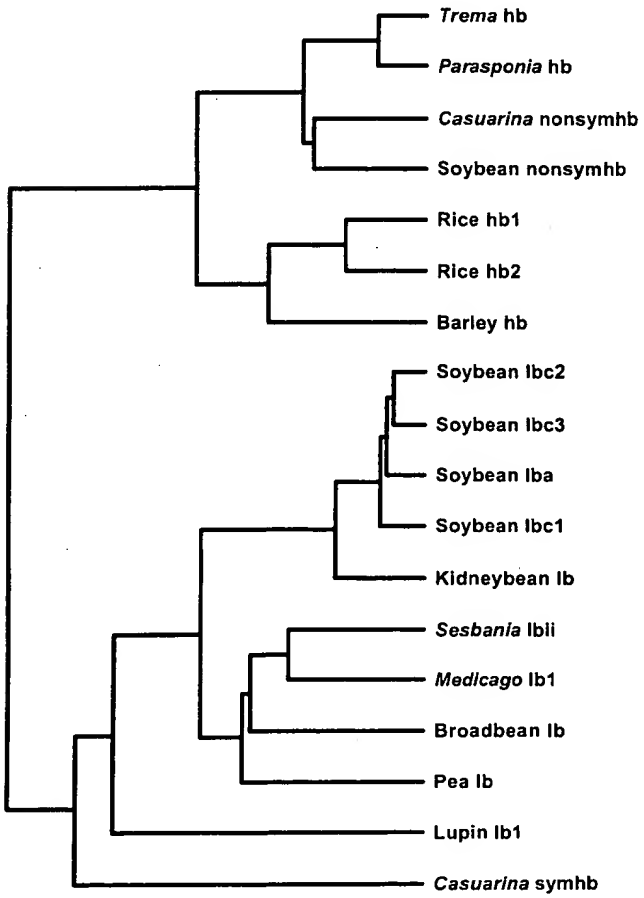

FIG. 4. Plant hemoglobin protein similarity tree. The similarity tree was constructed using the GCG PILEUP program. Sequences used are referenced in Fig. 1.

The Northern blot data are supported by *in situ* hybridization analyses that show that the soybean nonsymbiotic hemoglobin is expressed in both nodules and shoot tips. In nodules, the new hemoglobin gene is expressed in many cell types whereas leghemoglobins are present only in the bacteroid-containing cells (25).

DISCUSSION

The New Hemoglobin Gene Is Present in Soybean and Other Legumes and Encodes a "Nonsymbiotic" Protein. We have isolated a novel, expressed hemoglobin gene from soybean that is not a member of the leghemoglobin gene family. The gene has the characteristic four exon, three intron structure of all known plant hemoglobins. The predicted amino acid sequence is only 57–58% identical to leghemoglobin sequences (Table 1), but contains the critical heme and ligand binding residues of a functional hemoglobin. It resembles other nonsymbiotic hemoglobins in having an extension of five amino acids at the amino terminus (Fig. 1, Table 1). This new soybean gene is most abundantly expressed in stems and cotyledons of seedlings, and in roots, some nodule tissue, and young leaves of mature plants. Expression of the new soybean hemoglobin therefore differs considerably from that of the nodule-specific leghemoglobins and resembles that of the nonsymbiotic gene of *C. glauca* (7). Both the gene sequence and expression pattern indicate that we have isolated a nonsymbiotic soybean hemoglobin gene. Cross-hybridization of this soybean gene with sequences in the genomes of other legumes suggests that all legumes contain a similar nonsymbiotic hemoglobin in addition to their well-characterized leghemoglobin genes.

Evolution of Plant Hemoglobins. The isolation of this gene from soybean adds another important example to a growing list of plant hemoglobins whose expression is not associated with any symbiotic interactions with micro-organisms. The GenBank EMBL data base contains the 5′ sequences of two classes of rice expressed sequence tags with extensive homology to the nonsymbiotic hemoglobins (Fig. 1). The presence of expressed nonsymbiotic hemoglobin genes in a number of dicots, now including the legumes, as well as the monocots

```
soyhb          -184  TAAGCCACACAAATGGGAATGA-CTCCCCATTACA-ATGAAGGGCCAA--CTTTCATTTTCAATGAA-----TCCCACTATAAA  -110

Parahb         -109  TAAAAAACCCAAAGAT--ATGG-CTCCCCAATACCCT-GAAGAGTTACACACGA--------TCCCCATTTTTTCTACTATATA  -38
Tremahb        -117  AAAAAAACCCAAGGAG--ATGG-CTCTCCAGTACCCT-GAAGAGTTACATCTTATGATCTTATCCCCCATTTTTTCCATATATA  -38
Casnonsymhb    -208  CAATTGACCCAAAGAA--ATGG-CTTTC---GACCCACGAAGAGCCGGAGCTATCCCTGTTACG---------TGCGCTATATA  -140

soylbc3        -141  AAGTTTTTGAAAAGATGA-TTGTCTCTTCACCATACCA---ATTGAT-------CACCCTCCTCCAACA---AGCCAAGAGAGAC  -71
soylba         -133  AAATTTTTTAAAAGATCG-TTGTTTCTTCTTCATCATGCTGATTGA--------CACCCTCC--------ACAAGCCAAGAGAAAC  -65
soylbc1        -128  AGGATTTTGAAAAGATCA-TTGGCTCTTCGTCATGCCG---ATTGA--------CACCCTCC--------ACAAGCCAAGAGAAAC  -63
soylbc2        -139  AAGATTTTGAAAAGATCATTTGGCTCTTCATCATGCCG---ATTGA--------CACCCTCC--------ACAAGCCAAGAGAAAC  -63
Cassymhb1      -490  ACTTCAATCCCAAGATGTCCTCTCTCTTATTGATATTTGAACAACAACAAAGATAAACAACCATTATCCCTACCAAGCAAGTAA  -408
Cassymhb1      -324  ACTTCAATCCCAAGATGTCCTCTCTCTTATTGATATTTGAACAACAACAAAGATAAACAACCTTTATCCTTACCAAGCAGGTAA  -241
```

FIG. 5. Sequence alignment of the promoter regions of hemoglobin genes from soybean, *Parasponia, Trema*, and *Casuarina*. Gaps, indicated by dashes, have been introduced to maximize similarity. The nodulin motifs, 5′-AAAGAT and 5′-CTCTT, are indicated in bold type, as is the putative nonsymbiotic motif 5′-GAAGAG. The TATA box is underlined. *Casuarina glauca* symbiotic hemoglobin 1 (7) and *C. glauca* hb nonsymbiotic hemoglobin (6) and soybean nonsymbiotic hemoglobin (this study) are numbered from the start codon, *Parasponia andersonii* hemoglobin (3), *Trema tomentosa* hemoglobin (4), Soybean *lba*, *lbc1*, *lbc2*, *lbc3* (27) are numbered from the transcription start. The *C. glauca* symbiotic hemoglobin promoter contains a direct repeat in this region and both copies are included.

barley, wheat, maize (8), and rice, shows that nonsymbiotic hemoglobin is widespread and possibly ubiquitous in the plant kingdom. It is likely that the more specialized symbiotic hemoglobins arose by gene duplication of a preexisting "nonsymbiotic" gene. The presence in the ancient tetraploid soybean of two gene clusters of leghemoglobins (26), and now of two nonsymbiotic hemoglobin genes (Fig. 2), is consistent with an origin for the leghemoglobins by a gene duplication event occurring before the hybridization of the diploid progenitors of *G. max*.

A similarity-tree based on pairwise comparisons of many of the known plant hemoglobin amino acid sequences was constructed using the Genetics Computer Group (GCG; Madison, WI) PILEUP program (Fig. 4). All the nonsymbiotic hemoglobins cluster together, including the newly characterized gene from soybean. Within the nonsymbiotic hemoglobins, the dicot and monocot sequences form separate subclusters. All of the leghemoglobin sequences cluster together, with those from plants with determinate, indeterminate and lupin-type nodules forming separate subgroups, consistent with the analysis of Marcker and Sandal (26). The symbiotic hemoglobins of one species are more similar to symbiotic hemoglobins of other species than to the nonsymbiotic hemoglobins within the same species. This analysis suggests that leghemoglobins arose before speciation within the legumes, the result of an initial gene duplication event that was followed by further duplications giving rise to the symbiotic leghemoglobin gene families within the present day legumes.

The *C. glauca* symbiotic hemoglobin (cassymhb) clusters with the leghemoglobins rather than with the nonsymbiotic hemoglobins. The *C. glauca* nonsymbiotic hemoglobin protein is 87% similar to the soybean nonsymbiotic hemoglobin and groups with it on a separate subbranch of the dendrogram (Fig. 4). The promoter motifs of these two nonsymbiotic genes are also similar (Fig. 5). Because the symbiotic hemoglobins (cassymhb and the leghemoglobins) are more similar to each other than to their putative progenitor nonsymbiotic hemoglobins in sequence, expression and regulation, they may have been derived from the same initial gene duplication before the divergence of *Casuarina* and the legumes. The similarities of the symbiotic and nonsymbiotic hemoglobins from *Casuarina* and soybean suggests that the species are more closely related than has previously been thought (28). A study based on the DNA sequences of the large subunit of the Rubisco gene has also placed the legumes close to *Casuarina* (29).

Because of the animal/plant cross-kingdom occurrence of hemoglobin, the conserved functional amino acid residues in the proteins and the conserved positions of the introns within the genes, hemoglobins must be very ancient. They presumably predate the divergence of plants and animals and existed well before the diversification of the Angiosperms. We predict that hemoglobin genes will also be present in the genomes of more primitive plant groups, such as bryophytes, pteridophytes, cycads, and gymnosperms.

Possible Functions of Nonsymbiotic Plant Hemoglobin. Two possible functions of nonsymbiotic plant hemoglobins have been proposed, as a sensor of oxygen concentration, or as a carrier in oxygen transport (9). Some hemoglobins are known to regulate associated enzymatic activities by changes in their conformation due to substrate binding (30, 31). Both the *C. glauca* and soybean nonsymbiotic hemoglobin genes have moderate levels of expression, and the proteins may therefore be more abundant than is expected for an oxygen sensor. The induction in barley of hemoglobin expression by low oxygen tension (8) suggests that hemoglobin may be a normal component of the anaerobic response in plants, presumably to facilitate oxygen diffusion at low oxygen concentrations and supports a role for nonsymbiotic hemoglobin in oxygen transport.

Nonsymbiotic hemoglobin may facilitate oxygen diffusion in rapidly dividing cells, such as those in the root meristem. Our Northern blot analyses show that expression of the soybean nonsymbiotic hemoglobin gene is higher in the root elongation zone than in the root tip, suggesting that any requirement for oxygen transport is not confined to meristematic cells. The high levels of expression in cotyledons and stems also suggests that nonsymbiotic hemoglobin gene expression does not correlate with cell division, but is perhaps associated with high levels of metabolic activity. Nonsymbiotic hemoglobin genes are all expressed in various metabolically active tissues such as developing seeds and roots (4, 8) or in the vascular tissues of leaves, stems and seedling cotyledons (ref. 8 and this study). These are all sites of intense short distance solute transfer, an energy demanding process. It is possible that the nonsymbiotic hemoglobin is facilitating intracellular diffusion of oxygen to the mitochondria in metabolically active cells in order to meet an increased demand for oxidative respiration.

The widespread distribution and expression of the nonsymbiotic hemoglobin genes implies that they play an important role(s) in the metabolic biochemistry of all plants, perhaps comparable with the oxygen binding roles of hemoglobins in animal systems.

We thank Janice Norman for excellent technical assistance, Fraser Bergersen and Inara Licis for providing soybean plants, Robert Hill for making information regarding the barley hemoglobin gene available before publication, and U. H. Jacobsen for providing unpublished *in situ* hybridization data showing the expression of the soybean nonsymbiotic hemoglobin gene in a number of cell types in soybean nodules.

1. Appleby, C. A. (1984) *Annu. Rev. Plant Physiol.* **35**, 443–478.
2. Appleby, C. A., Tjepkema, J. D. & Trinick, M. J. (1983) *Science* **220**, 951–953.

3. Landsmann, J., Dennis, E. S., Higgins, T. J. V., Appleby, C. A., Kortt, A. A. & Peacock, W. J. (1986) *Nature (London)* **324,** 166–168.
4. Bogusz, D., Appleby, C. A., Landsmann, J., Dennis, E. S., Trinick, M. J. & Peacock, W. J. (1988) *Nature (London)* **331,** 178–180.
5. Fleming, A. I., Wittenberg, J. B., Wittenberg, B. A., Dudman, W. F. & Appleby, C. A. (1987) *Biochim. Biophys. Acta* **911,** 209–220.
6. Christensen, T., Dennis, E. S., Peacock, J. W., Landsmann, J. & Marcker, K. A. (1991) *Plant Mol. Biol.* **16,** 339–344.
7. Jacobsen-Lyon, K., Jensen, E. O., Jorgensen, J.-E., Marcker, K. A., Dennis, E. S. & Peacock, W. J. (1995) *Plant Cell.* **7,** 213–223.
8. Taylor, E. R., Nie, X. Z., MacGregor, A. W. & Hill, R. D. (1994) *Plant Mol. Biol.* **24,** 853–862.
9. Appleby, C. A., Bogusz, D., Dennis, E. S. & Peacock, W. J. (1988) *Plant Cell Environ.* **11,** 359–367.
10. Miao, G.-H. & Verma, D. P. S. (1993) *Plant Cell* **5,** 781–794.
11. Hirano, H. (1989) *J. Protein Chem.* **8,** 115–130.
12. Wiborg, O., Hyldig-Nielsen, J. J., Jensen, E. O., Paludan, K. & Marcker, K. A. (1983) *EMBO J.* **2,** 449–452.
13. Davis, L. G., Dibner, M. D. & Battey, J. F. (1988) *Basic Methods in Molecular Biology* (Elsevier, New York).
14. Burke, T. J., Callis, J. & Vierstra, R. D. (1988) *Mol. Gen. Genet.* **213,** 435–443.
15. Konieczny, A. (1987) *Nucleic Acids Res.* **15,** 6742.
16. Hyldig-Nielsen, J. J., Jensen, E. L., Paludan, K., Wiborg, O., Garrett, R., Jorgensen, P. & Marcker, K. A. (1982) *Nucleic Acids Res.* **10,** 689–701.
17. Wiborg, O., Hyldig-Nielsen, J. J., Jensen, E. O., Paludan, K. & Marcker, K. A. (1982) *Nucleic Acids Res.* **10,** 3487–3494.
18. Lee, J. S. & Verma, D. P. S. (1984) *EMBO J.* **3,** 2745–2752.
19. Metz, B. A., Welters, P., Hoffman, H. J., Jensen, E. O., Schell, J. & de Bruijn, F. J. (1988) *Mol. Gen. Genet.* **214,** 181–191.
20. Gallusci, P., Dedieu, A., Journet, E. P., Huguet, T. & Barker, D. G. (1991) *Plant Mol. Biol.* **17,** 335–349.
21. Richardson, M., Dilworth, M. J. & Scawen, M. D. (1975) *FEBS Lett.* **51,** 33–37.
22. Lehtovaara, P., Lappalainen, A. & Ellfolk, N. (1980) *Biochim. Biophys. Acta* **623,** 98–106.
23. Ramlov, K. B., Laursen, N. G., Stougaard, J. & Marker, K. A. (1993) *Plant J.* **4,** 577–580.
24. Szczyglowski, K., Szabados, L., Fujimoto, S. Y., Silver, D. & de Bruijn, F. J. (1993) *Plant Cell* **6,** 317–332.
25. Nap, J.-P. & Bisseling, T. (1990) *Science* **250,** 948–954.
26. Marcker, K. A. & Sandal, N. N. (1987) in *Plant Molecular Biology,* eds. von Wettstein, D. & Chua, N.-H. (Plenum, New York), pp. 503–508.
27. Stougaard, J., Sandal, N. N., Gron, A., Kuhle, A. & Marcker, K. A. (1987) *EMBO J.* **6,** 3565–3569.
28. Cronquist, A. (1981) *An Integrated System of Classification of Flowering Plants* (Columbia Univ. Press, New York).
29. Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., *et al.* (1993) *Ann. Missouri Bot. Gard.* **80,** 528–580.
30. Zhu, H. & Riggs, A. F. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 5015–5019.
31. Sherman, D. R., Guinn, B., Perdok, M. M. & Goldberg, D. E. (1992) *Science* **258,** 1930–1932.

The Plant Cell, Vol. 7, 213–223, February 1995 © 1995 American Society of Plant Physiologists

Symbiotic and Nonsymbiotic Hemoglobin Genes of *Casuarina glauca*

Karin Jacobsen-Lyon,[a,b,1] Erik Østergaard Jensen,[b] Jan-Elo Jørgensen,[b] Kjeld A. Marcker,[b] W. James Peacock,[a] and Elizabeth S. Dennis[a,2]

[a] Commonwealth Scientific and Industrial Research Organization, Division of Plant Industry, GPO Box 1600, Canberra ACT 2601, Australia

[b] Laboratory of Gene Expression, Department of Molecular Biology, Aarhus University, Gustav Wiedsvej 10, DK-8000 Aarhus, Denmark

***Casuarina glauca* has a gene encoding hemoglobin (*cashb-nonsym*). This gene is expressed in a number of plant tissues. *Casuarina* also has a second family of hemoglobin genes (*cashb-sym*) expressed at a high level in the nodules that *Casuarina* forms in a nitrogen-fixing symbiosis with the actinomycete *Frankia*. Both the nonsymbiotic and symbiotic genes retained their specific patterns of expression when introduced into the legume *Lotus corniculatus*. We interpret this finding to mean that the controls of expression of the symbiotic gene in *Casuarina* must be similar to the controls of expression of the leghemoglobin genes that operate in nodules formed during the interaction between rhizobia and legumes. Deletion analyses of the promoters of the *Casuarina* symbiotic genes delineated a region that contains nodulin motifs identified in legumes; this region is critical for the controlled expression of the *Casuarina* gene. The finding that the nonsymbiotic *Casuarina* gene is also correctly expressed in *L. corniculatus* suggests to us that a comparable nonsymbiotic hemoglobin gene will be found in legume species.**

INTRODUCTION

Nitrogen-fixing symbioses between plants and bacteria have been studied in detail in legumes, but a number of other plant families also have symbiotic nitrogen-fixing associations. For example, the woody shrub and tree species of *Casuarina* fix nitrogen in association with the actinomycete *Frankia*. Just as in the nodules of legumes, hemoglobin is present in the *Casuarina* nodules, where it transports oxygen to the symbiont (Fleming et al., 1987).

Kortt et al. (1988) showed that *Casuarina* nodule hemoglobin has 44% amino acid sequence identity with the leghemoglobin of soybean nodules. We isolated a hemoglobin gene from *C. glauca* using a *Parasponia* hemoglobin gene probe (Landsmann et al., 1986). Contrary to our expectation, the deduced amino acid sequence of the *Casuarina* gene did not correspond to the amino acid sequence of the *Casuarina* nodule hemoglobin (Christensen et al., 1991). The gene encoded a protein with a predicted amino acid sequence having only 53% identity with the nodule protein, whereas it was 80% identical to *Parasponia* hemoglobin. In *Parasponia*, the hemoglobin gene is expressed at a high level in nodules and at a lower level in normal root tissues (Landsmann et al., 1986). These findings suggested to us that the *Casuarina* gene we had isolated was likely to be expressed in plant tissue other than nodules and that *Casuarina* must have another gene that is expressed in nodules. In *Parasponia*, one gene performs both of these roles.

In this study, we report the isolation of hemoglobin genes from *Casuarina*. These genes encode hemoglobin found in *Casuarina* nodules. We show that this hemoglobin gene family has a pattern of expression different from that of the previously isolated *Casuarina* gene. The newly isolated symbiotic gene family is expressed only in the nodule. Nodule-specific expression of leghemoglobin genes has been studied extensively in transgenic legumes. The promoter regions of the soybean *lbc3* and the *Sesbania rostrata glb3* genes both confer nodule-specific expression in *Lotus corniculatus* (Stougaard et al., 1986; Szabados et al., 1990). Promoter deletion and fusion analyses identified a *cis* regulatory element important for high-level nodule-specific expression (Stougaard et al., 1987; Szabados et al., 1990). This element (OSE [organ-specific element] in the *lbc3* promoter and NICE [nodule-infected cell expression] in the *glb3* promoter) contains two motifs, 5′-AAAGAT-3′ and 5′-CTCTT-3′, present in most nodulin promoters (Sandal et al., 1987). Site-directed mutagenesis showed that the contribution of the 5′-AAAGAT-3′ motif to nodule-specific expression is of minor importance, whereas the 5′-CTCTT-3′ motif as well as the adjacent 5′-TGG-3′ are essential

[1] Current address: Children's Medical Research Institute, Locked Bag 23, Wentworthville NSW 2145, Australia.

[2] To whom correspondence should be addressed.

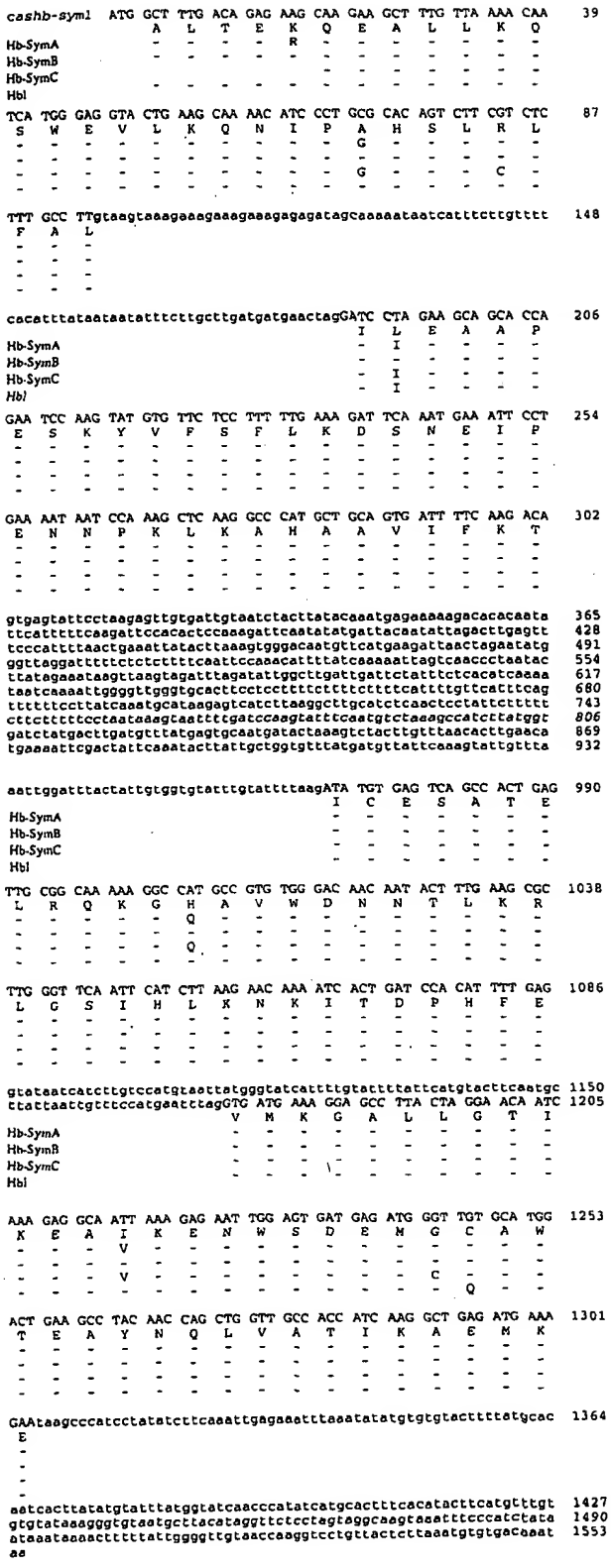

Figure 1. Sequence of the *Casuarina* Symbiotic Hemoglobin Gene (*cashb-sym1*).

for expression of the nodule-specific gene (Ramlov et al., 1993; Szczyglowski et al., 1994).

We introduced both the *Casuarina* symbiotic and nonsymbiotic hemoglobin promoters linked to the β-glucuronidase (*GUS*) reporter gene into *L. corniculatus*. The patterns of expression observed in *Casuarina* for both genes were retained in the transgenic legume. This suggests that in the two plant families not only do similar molecular mechanisms control the activity of the nodule-specific genes, but also common controls must exist for the plant tissue–expressed gene. This raises the possibility that legumes have a second family of hemoglobin genes that are expressed in plant tissues. These genes are yet to be described.

The promoter of the *Casuarina* symbiotic gene contains nodulin motifs similar to those found in leghemoglobin genes. Deletion analysis showed that the region containing the nodulin motifs was essential for expression of the reporter gene construct in transgenic nodules.

RESULTS

Casuarina Has Both Symbiotic and Nonsymbiotic *Hemoglobin Genes*

Hemoglobin cDNA clones were isolated from a *Casuarina* nodule expression library in screens with a polyclonal antibody raised against *Casuarina* nodule hemoglobin (Fleming et al., 1987). The sequence of one of the clones (Hb-SymA; Figure 1) has a deduced amino acid sequence almost identical to the protein sequence reported by Kortt et al. (1988) for purified nodule hemoglobin. Two other clones (Hb-SymB and Hb-SymC) had sequences not completely identical to Hb-SymA but showed at least 97% identity in the coding region (Figure 1). These could be other members of a gene family or they could be allelic forms of the one gene. The "gene family" alternative is supported by DNA gel blot analyses in which an Hb-SymA probe revealed a complex pattern of bands (Figures 2A and 2B). BamHI digests produced nine hybridizing bands, four strong bands (2.2, 4, 6.5, and 9.5 kb), and five weakly hybridizing bands (3.9, 8, 12, 15, and 19 kb). We determined that these bands represent alleles of a smaller number of genes rather than nine different genes by analyzing DNA from individual seedlings (Figure 2B). Four of nine individual seedlings (1, 2, 5, and 9) displayed a banding pattern identical to that of the pooled population, giving support to the multiple gene alternative. The remaining five plants had only slightly different hybridization patterns, suggesting some sequence polymorphism in the population. Only one of the strongly

The deduced amino acid sequences of three cDNAs (Hb-SymA, Hb-SymB, and Hb-SymC) are compared with the amino acid sequence of the nodule hemoglobin Hbl (Fleming et al., 1987). Lowercase letters indicate the noncoding sequence. Dashes indicate identity with the amino acid encoded by *cashb-sym1*.

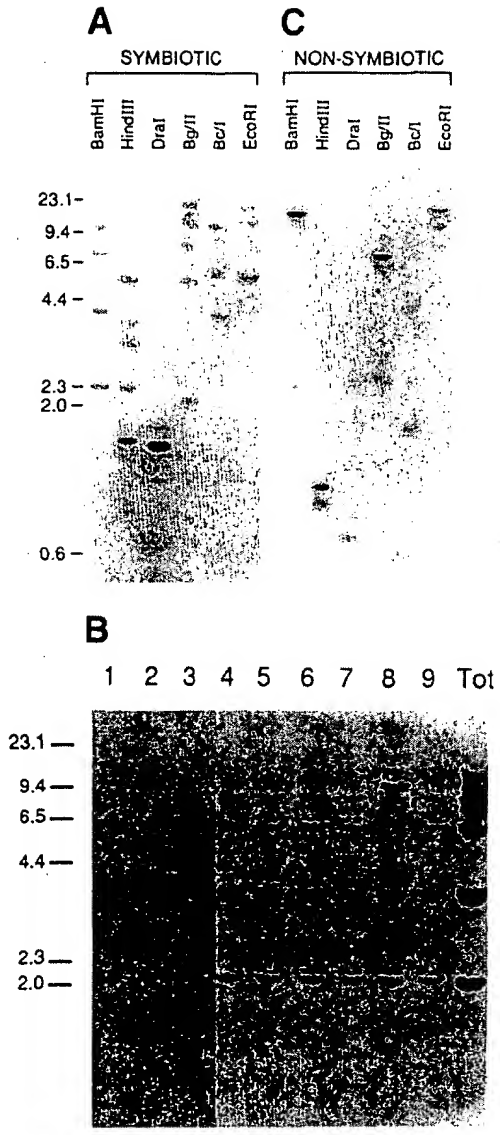

Figure 2. DNA Gel Blot Analysis of *C. glauca* Genomic DNA.

(A) Restriction digest hybridized to Hb-SymA cDNA is shown.
(B) BamHI digests of nine individual seedlings (labeled above the gel) probed with Hb-SymA cDNA are shown.
(C) Restriction digests hybridized to a fragment of the *cashb-nonsym* gene are shown.
Molecular length markers are given at left in kilobases. Tot, total population of plants.

hybridizing bands showed length polymorphism; in plant 8, the 6.5-kb band was missing but a new, strong band was seen at 11 to 12 kb. The four strongly hybridizing bands, which must represent sequences with high homology to the cDNA probe, are therefore likely to be from a set of genes present in each of the plants. The less intense bands could represent inactive pseudogenes with lower homology. Thus, *Casuarina* resembles legumes that also have three or four symbiotic leghemoglobin genes.

Previously, we had isolated a hemoglobin gene from *C. glauca* using a *Parasponia* hemoglobin gene as a probe (Christensen et al., 1991). This *Casuarina* gene showed a low level of deduced amino acid similarity with the nodule protein. DNA gel blot analysis with this gene revealed a simple genomic pattern, with most restriction enzymes generating only one strongly hybridizing band and one weakly hybridizing band (Figure 2C). During the isolation of this gene, we had detected, in addition to the full-length gene, a fragment containing an incomplete gene sequence (Christensen et al., 1991); this fragment probably corresponds to the weaker band (Figure 2C). We concluded that *C. glauca* has only one functional nonsymbiotic hemoglobin gene and, probably, a truncated nonfunctional gene segment.

We isolated 30 clones from a genomic library of 240,000 plaque-forming units by using Hb-SymC as a probe. Seven of the clones mapped into three classes (Figure 3). Classes 1 and 2 contained only a single gene, whereas class 3 clones (1, 18, and 21) each contained two hemoglobin genes in opposite polarity, 7 kb apart.

The sequence of one of the genes, *cashb-sym1* (genomic clone 38) (Figure 1), has a deduced amino acid sequence identical to that of Hb-SymB, implying that it is transcribed in nodules. The gene has three introns in exactly the same positions as the introns of all known plant hemoglobin genes (Jensen et al., 1981; Landsmann et al., 1986), including the *Casuarina* nonsymbiotic hemoglobin gene (Christensen et al., 1991).

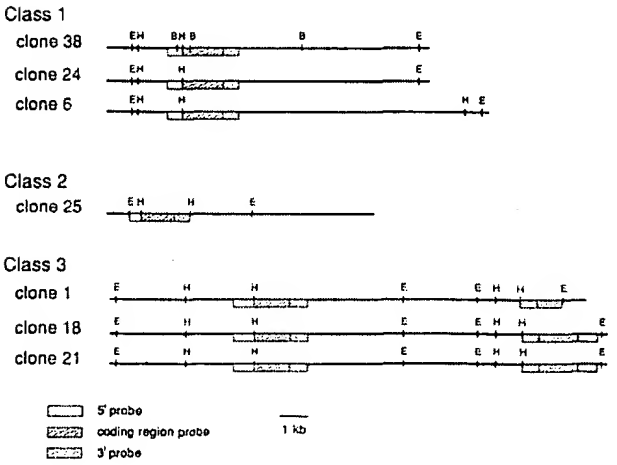

Figure 3. Diagram of Organization of Genomic Hemoglobin Segments Cloned from *Casuarina*.

Seven clones were characterized, and these fall into three classes. Class 3 has two copies of the gene on each clone. The various probes used to map the genes are shown. B, E, and H are BamHI, EcoRI, and HindIII restriction enzyme sites, respectively.

The Symbiotic and Nonsymbiotic Genes Have Different Expression Patterns

RNA gel blot hybridization using the Hb-SymA probe on RNA isolated from a number of *Casuarina* tissues showed that the symbiotic genes are expressed at a high level in nodules. No expression was detectable in root, leaf, or stem tissue (Figure 4A). The nonsymbiotic gene was expressed in root, leaf, and stem tissues and only at a barely detectable level in nodules (Figure 4B). The level of expression in the nodule was ~100-fold less than that of the symbiotic genes. The hemoglobin produced by the symbiotic gene must play a role in the *Frankia* symbiosis; this role is probably comparable to the oxygen transport role of leghemoglobin in rhizobia symbioses with legumes. In contrast, the nonsymbiotic hemoglobin protein is unlikely to be involved in nitrogen fixation, but rather to be involved in other metabolic processes in a variety of plant tissues.

Casuarina Hemoglobin Genes Maintain Their Expression Patterns in Transgenic *L. corniculatus*

The promoters of members of the two gene families were fused to a *GUS* reporter gene cassette and introduced into *L. corniculatus*, a legume species (Stougaard et al., 1986). The constructs involving the promoters of two *cashb-sym* genes showed a high level of expression in the central, infected cells of the nodule (Figure 5A). The level of GUS activity generated by the promoter of the *Casuarina* symbiotic gene in the infected cells of the nodule is similar to that seen with the soybean *lbc3* leghemoglobin promoter in transgenic *L. corniculatus* (Lauridsen et al., 1993). By using dark-field microscopy, a technique that increases the sensitivity of detection of *GUS* reporter gene activity (Medberry et al., 1992), we were able to detect low-level expression in the cortical cells lying just outside the central infected zone (see pink cells, Figure 5B), but there was no expression detectable in the other cells of the cortex or in any of the cells of the vascular traces. Dark-field microscopy also showed low-level expression in some of the interstitial uninfected cells in the central zone of the nodule (Figure 5B). In addition to expression in the nodules of the transgenic roots, the *cashb-sym* promoter was active in the root cap cells of roots that had been cultured without exposure to *Rhizobium* (Figure 5H).

The promoter of the nonsymbiotic gene also directed activity of the *GUS* reporter gene in nodules, but only at a level detectable by the sensitive dark-field technique (Figure 5D) and not in bright-field view (Figure 5C). Expression occurred in the nodule parenchyma cells between the endodermis and the infected central cells and in the vascular bundles of the nodule (Figure 5D). This pattern overlaps the expression pattern of the symbiotic genes, which are also expressed in cells immediately surrounding the central zone. Dark-field analysis showed that the nonsymbiotic gene was also expressed in the uninfected but not the infected cells of the central zone (Figure 5D). The blue color of the infected cells in Figure 5D is not due to GUS staining. The gene was active in the meristematic zone of root tips (Figure 5G) and in the parenchyma internal to the endodermis and associated with the vascular stele of the root (cf. Figures 5E and 5F).

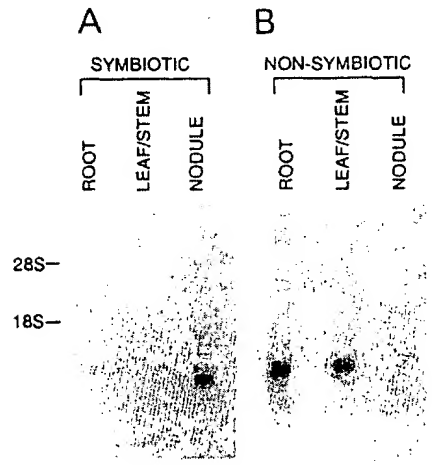

Figure 4. The Expression Pattern of *Casuarina* Hemoglobin Genes.

(A) Gel blot analysis of RNA from various *Casuarina* tissues using a *cashb-sym* gene probe.

(B) A *cashb-nonsym* gene was used to probe the same blot shown in **(A)**. Equal amounts of total RNA, as determined at $OD_{260\,nm}$, from nodules, roots, and leaves/stems were hybridized. 28S and 18S, rRNAs.

The *cashb-sym* Promoters Have Nodulin Motifs

In legumes, the promoters of hemoglobin genes and other nodulin genes contain two motifs, 5′-AAAGAT-3′ and 5′-CTCTT-3′, that are 6 bp apart in most cases (Figure 6). The sequences are critical for the nodule-specific expression of the genes (Stougaard et al., 1987; Ramlov et al., 1993; Szczyglowski et al., 1994). The promoters of the *cashb-sym* genes also contain these same two nodulin motifs in similar relative positions (Figures 6 and 7). This suggests that these motifs have the same function in the hemoglobin genes of the *Casuarina–Frankia* symbiosis as they do in genes of the legume–rhizobia symbiosis. The regions surrounding the *Casuarina* nodulin boxes are AT rich and show some homology with the comparable regions of the legume promoters (Figure 6). There is a second copy of the AAAGAT motif farther downstream in both *cashb-sym* promoters (Figure 7).

We examined the functional importance of the nodulin motifs in the *Casuarina* genes by 5′ deletion analysis (Figure 8). Removal of the nodulin motif region (−252-bp deletion) disabled the *cashb-sym1* promoter (Figure 8A). In this particular gene, there is a duplication of the segment including the nodulin motifs (−478 to −462 cf. −311 to −295; Figure 7); deletion of the proximal motifs (Δ344/186; Figures 8A and 8B) showed

that the remaining upstream motif sequence was not sufficient to promote gene activity in the nodule. If the distal motifs were disrupted (−474; Figure 8A), expression levels were reduced but were still detectable and the expression pattern was unchanged. The *cashb-sym2* gene has only one copy of each of the two motifs, although it has levels of expression comparable to those of the *cashb-sym1* gene (Figures 8A and 8B). The deletion analysis also showed that other upstream regions were important for high-level expression (Figure 8B). The removal of these regions did not alter specificity of expression but did alter the level of expression.

The *cashb-nonsym* promoter contains sequences related but not identical to nodulin motifs (Figure 6), and the spacing (4 bp) between the motif sequences differs from that in the symbiotic genes (Figure 6). A 5′ deletion analysis showed that the promoter region containing the motifs does not control gene activity of this nonsymbiotic gene (Figures 8C and 8D). Sequences upstream of −1000 are important for both root and nodule expression.

DISCUSSION

Nitrogen-fixing symbioses between plants and bacteria occur in a number of dicotyledonous families. In a number of phylogenetic schemes, these nitrogen-fixing families do not seem to be closely related (Landsmann et al., 1986). The best characterized symbiosis is that between legumes and rhizobia. Only one non-legume symbiosis involving rhizobia occurs; this is in the genus *Parasponia* (Ulmaceae, the elm family). In other nitrogen-fixing families of plants, such as the Casuarinaceae and Myricaceae, the symbiotic relationship is with the Gram-positive actinomycete *Frankia*. The differing modes of invasion by the symbiotic organisms and the differing morphologies of the nodules formed in the symbioses have suggested to a number of investigators that the rhizobia- and *Frankia*-based symbioses have evolved independently (Appleby et al., 1988; Schwintzer and Tjepkema, 1990).

In both types of symbioses, the biochemistry is similar. Nitrogen fixed by the microsymbiont is exchanged for photosynthate produced by the host plant. Both symbiotic systems employ hemoglobin to supply oxygen needed for microbial respiration at free oxygen concentrations sufficiently low to protect the oxygen-sensitive bacterial nitrogenase enzyme systems (Appleby, 1984).

Initially, when plant hemoglobin was known only from legumes, it was thought that the gene may have been introduced into plants from the animal kingdom by some lateral evolutionary mechanism (Appleby, 1974; Jeffreys, 1981; Hyldig-Nielsen et al., 1982). Although the amino acid sequences of plant and animal hemoglobins have limited sequence identity (13 to 15%), the key functional residues in the molecules are conserved, and the two introns occurring in most animal hemoglobins are in precisely the same positions as two of the three introns found in all plant hemoglobins (Jensen et al., 1981; Landsmann et al., 1986; Christensen et al., 1991). The position of the central third intron in plants had been predicted by Go (1981) on the basis of structural analysis of the animal kingdom hemoglobin molecule. These data and the fact that hemoglobin genes are known in several unrelated families of plants make the lateral gene transfer hypothesis unlikely. Instead, our observations favor a single evolutionary origin of the hemoglobin genes of both the animal and plant kingdoms.

In our current studies, we have isolated the symbiotic hemoglobin genes of *Casuarina* and shown that the members of this gene family bear striking similarities with leghemoglobin genes. They have the same major features of gene organization and the same nodule-specific expression, which is restricted to the central bacteroid-containing region; they also have nodulin sequence motifs in their promoters. In both legumes and *Casuarina*, the nodulin motif region is essential for gene activity in nodules.

We have also found that when the *Casuarina* symbiotic genes are introduced into a legume (*L. corniculatus*), they have a high level of expression in the bacteroid-infected zone of the nodule, indicating that the *Casuarina* promoter motifs are recognized by the transcriptional signal systems of the legume, despite the different symbiotic microorganisms of the two systems. The expression in the infected cells may be dependent upon a signal produced by the invading bacteria, whether they are rhizobia or *Frankia*. De Billy et al. (1991) suggested the presence of such signals in legumes, and Welters et al. (1993) have identified a bacterial DNA binding protein interacting with the *S. rostrata* leghemoglobin promoter. We are not able to explain the significance of our observation of some gene activity in the root cap zone in noninfected transgenic roots (Figure 5H), although it could be a consequence of the *A. rhizogenes* transformation system; the transgenic hairy roots are known to have an altered hormone regime compared with normal roots (Schmülling et al., 1988).

The leghemoglobin and the *Casuarina* symbiotic genes contain nodulin motifs at comparable relative positions in their promoters. Similarly spaced motifs have also been recorded in other nodulin genes with expression restricted to nodules (Sandal et al., 1987; Stougaard et al., 1990). The nonsymbiotic hemoglobin gene of *Casuarina* does not have nodulin motifs closely matching the consensus motifs nor is their spacing comparable to that of the motifs of the symbiotic genes (Figure 6). This gene does not have a nodule-specific expression pattern. It is active in a number of plant tissues and in the nodule, where it is expressed only in the vascular and inner cortical regions. The structure of the nonsymbiotic gene of *Casuarina* is similar to that of leghemoglobin genes in that it has the same intron positions, but its deduced amino acid sequence more closely resembles that of the hemoglobin gene of the non-legume *Parasponia*, with 80% sequence identity and with an N-terminal extension to the protein sequence (Christensen et al., 1991) (Table 1), as is found in the *Parasponia* gene.

From existing data (Wittenberg et al., 1974; Fleming et al., 1987; Gibson et al., 1989), it is reasonable to assume that the

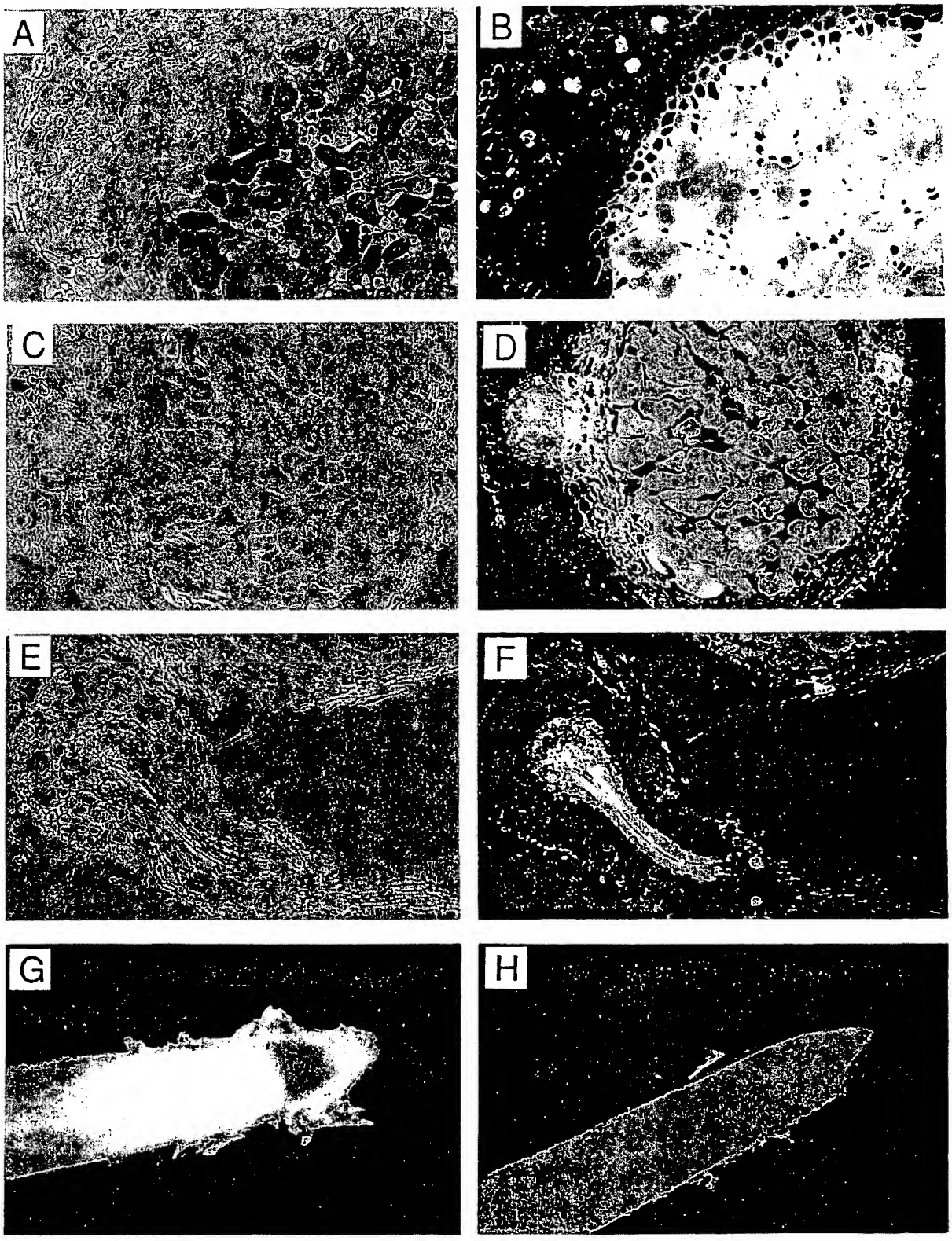

Figure 5. Expression of Chimeric *Casuarina* Hemoglobin Genes in *L. corniculatus*.

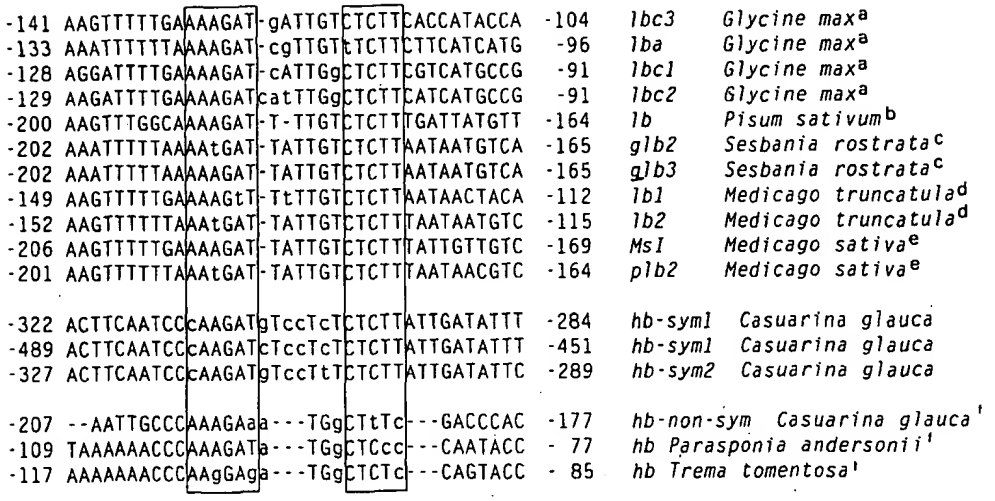

Figure 6. Alignment of Promoter Sequences from Different Symbiotic and Nonsymbiotic Hemoglobin Genes.

The nodulin motif sequences are boxed. Uppercase letters between and in the boxed sequences indicate identity with the consensus sequence determined for nodule genes. Numbers indicate the distance from the transcription initiation site, except for the *Casuarina* sequences, in which numbers indicate the distance from the translation initiator ATG. References indicated by superscript letters are as follows: a, Stougaard et al. (1987); b, Nap (1988); c, Metz et al. (1988); d, Gallusci et al. (1991); e, Davidowitz et al. (1991); f, Christensen et al. (1991). The *cashb-sym* sequences are from this study.

hemoglobins in the nodules of legumes, *Casuarina*, and *Parasponia* all function as oxygen carriers in the symbiotic interaction. The nonsymbiotic hemoglobin in *Casuarina* also has all the sequence and structural characteristics of an oxygen carrier. The fact that its expression is restricted to plant tissues where oxygen supply is likely to be limiting adds support to this possibility. Recently, it has been reported that a hemoglobin gene exists in barley and that it is induced by hypoxic conditions (Taylor et al., 1994), supporting an oxygen transport role. We had earlier suggested another possible function for hemoglobins in plants as oxygen sensors rather than oxygen carriers (Appleby et al., 1988). The oxygen sensor (Fix L protein) of *R. meliloti* has been shown to be a hemoprotein with kinase activity (Gilles-Gonzales et al., 1991). Until we have mutant plants available, we will not be certain of the function of plant hemoglobins in nonsymbiotic tissues.

Evolution of Hemoglobin Genes in Plants

The sequence and structural identities of *Casuarina* nonsymbiotic hemoglobin and *Parasponia* hemoglobin suggest a direct evolutionary lineage for these genes. Previously, we had suggested that it was likely that all plants have a hemoglobin gene and predicted that in addition to the symbiotic leghemoglobins known in legumes, there may be another gene or gene family in legumes encoding hemoglobins operative in nonsymbiotic plant tissues (Appleby et al., 1988; Bogusz et al., 1988). This concept has been given additional support by the findings of hemoglobin genes in barley and maize (Taylor et al., 1994).

In *Casuarina*, the nonsymbiotic hemoglobin gene could have given rise to the specialized symbiotic hemoglobin gene family by gene duplication and subsequent sequence divergence. In *Parasponia*, the requirement for high-level expression in

Figure 5. (continued).

Tissue was stained, and thin sections were cut and examined by bright- or dark-field microscopy. Expression of the *cashb-sym1* promoter–*GUS* fusion is shown in **(A)**, **(B)**, and **(H)**.

(A) Bright-field micrograph of a nodule cross-section showing GUS activity (blue) in the symbiont-infected cells.

(B) Dark-field micrograph of the same section shown in **(A)**. A high level of GUS activity is shown in the infected cells (blue), and a low level of GUS activity (pink) is visible in uninfected interstitial cells of the central zone and in the inner cortical cells immediately surrounding the central zone.

(C) to **(G)** Expression of the *cashb-nonsym* promoter–*GUS* fusion. **(C)** is a bright-field micrograph of a nodule cross-section showing no visible GUS activity. **(D)** is a dark-field micrograph of the same section shown in **(C)**. GUS activity (pink) is now visible in the central zone cells between the infected cells and in the inner cortex of the nodule parenchyma. The blue appearance of the bacteroids is not the result of *GUS* expression. A higher level of activity may be present in the nodule vascular tissues. **(E)** shows a cross-section through a root with a lateral root (bright-field). **(F)** is a dark-field micrograph of the same section shown in **(E)**. GUS activity (pink) occurs in the stelar parenchyma of the main root and lateral root; a higher level of activity may be present in the vascular cells. Expression in the meristematic zone of a root tip is shown in **(G)**.

(H) Expression of *cashb-sym1* in the root cap cells of a root not infected with *Rhizobium*.

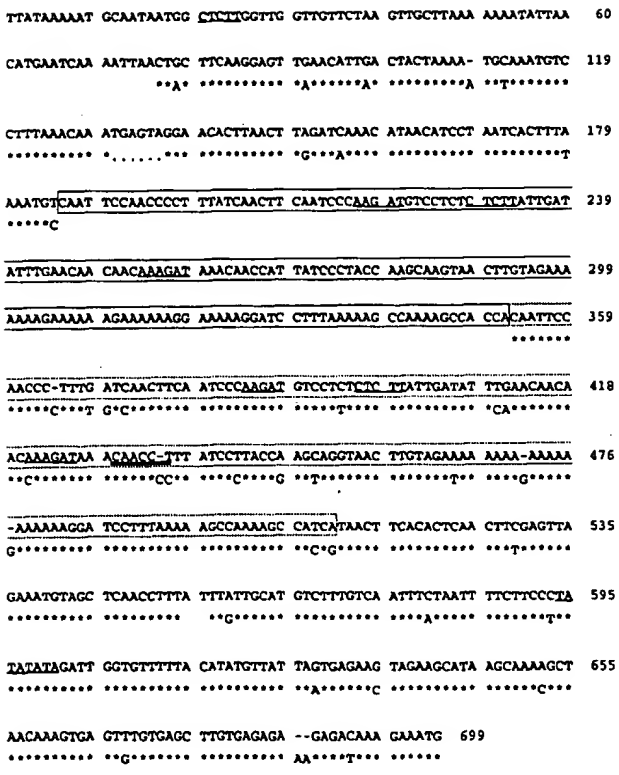

Figure 7. Sequence of the Promoter of *Casuarina* Symbiotic Hemoglobin Genes.

The sequence of the *cashb-sym1* promoter is aligned with that of the *cashb-sym2* promoter. An asterisk indicates an identical base; dashes were introduced to optimize alignment; dots indicate a deletion. The 158-bp duplication of *cashb-sym1* is indicated by boxes. The TATA box with nodulin motifs is underlined. The proposed CACCCT box is double underlined.

symbiotic nodules and lower expression in nonsymbiotic tissue has apparently been achieved with a single gene with differential expression in nodules and nonsymbiotic tissues of the plant.

Symbiotic hemoglobin genes could have arisen independently in rhizobia and *Frankia* symbioses, or alternatively, a gene duplication event could have happened before the divergence of the legume and *Casuarina* plant families. The hypothesis of independent origins of the symbiotic genes in the two families has some support in the observation that *Casuarina* symbiotic hemoglobin shows more sequence similarity with *Casuarina* nonsymbiotic hemoglobin than with soybean symbiotic hemoglobins (53 versus 44% identity; Table 1). However, two lines of evidence favor a common origin of the symbiotic genes in legumes and *Casuarina*. First, the transcription factors in the *L. corniculatus* nodule are compatible with the *Casuarina* promoter sequences, and comparable nodulin motifs are critical in both leghemoglobin and in the *cashb-sym* genes. Second, both symbiotic genes lack the N-terminal extension found in all known nonsymbiotic genes.

The recruitment of preexisting genes for nodule-specific function by gene duplication is likely to have occurred with genes other than the hemoglobin gene. Glutamine synthase, glutamate synthase, the peribacteroid membrane proteins N-23, N-24, and N-26, and early nodulins all show differently regulated genes with either nodule-specific or non–nodule-specific expression patterns (Nap and Bisseling, 1990; Miao and Verma, 1993). This contrasts to the situation in *Parasponia*, in which a single gene is expressed differently in nodule and host plant tissues. It may be that *Parasponia*, the legumes, and *Casuarina* share a common ancestor that was involved in the early evolution of symbiotic nitrogen fixation and that subsequently there have been at least two different strategies of specialization for genes involved in nodule function in families of flowering plants.

METHODS

RNA/DNA Isolation Method from *Casuarina*

RNA was prepared from *Casuarina glauca* nodules by a modification of the method of Hughes and Galau (1988). Plant tissue (5 to 10 g) was ground to a fine powder in liquid nitrogen and sprinkled into 55 mL of cold buffer (one-tenth volume of TE3D [200 mM Tris-HCl, pH 8.5, 300 mM LiCl, 10 mM Na_2EDTA, 1.5% lithium dodecylsulfate, 1% w/v sodium deoxycholate, 1% v/v Nonidet P-40], 5% w/v insoluble PVP, 90 mM mercaptoethanol, 10 mM DTT, 0.1% DEPC) and stirred to ensure immediate contact with the buffer. After stirring for 5 to 10 min, 46 mL of 3 M ammonium acetate was added, and the extract was spun at 5000*g* for 20 min. RNA/DNA was precipitated from the supernatant with one-tenth volume of 3 M sodium acetate and one-half final volume of isopropanol and was spun at 5000*g* for 30 min. The pellet was resuspended in 5 to 10 mL of H_2O and purified by phenol–chloroform extraction; this process was repeated until the preparation looked clean. The preparation was finished with a chloroform extraction. RNA was precipitated with one-quarter volume of 10 M LiCl on ice for 2 to 12 hr, followed by a spin at 10,000*g* for 30 min. DNA was recovered from the supernatant by ethanol precipitation. The RNA pellet was resuspended, ethanol precipitated, and washed in 70% ethanol.

Isolation of cDNAs

Poly(A)$^+$ RNA was fractionated from nodule RNA. A nodule cDNA library in λZAPII (Stratagene) was constructed from 5 μg of poly(A)$^+$ RNA (cDNA synthesis kit; Pharmacia). One hundred and twenty plaque-forming units were screened using antibodies raised against *Casuarina* hemoglobin isolated from nodules (Fleming et al., 1987). Three positive clones were isolated and sequenced using Applied Biosystems (Foster City, CA) automatic sequencing system and standard sequencing procedures. The EMBL accession numbers for Hb-SymA, Hb-SymB, and Hb-SymC are X77694, X77695, and X77696, respectively.

DNA Gel Blot Analysis

Genomic DNA was isolated from *Casuarina* seedlings essentially as described by Hughes and Galau (1988) and digested with enzymes

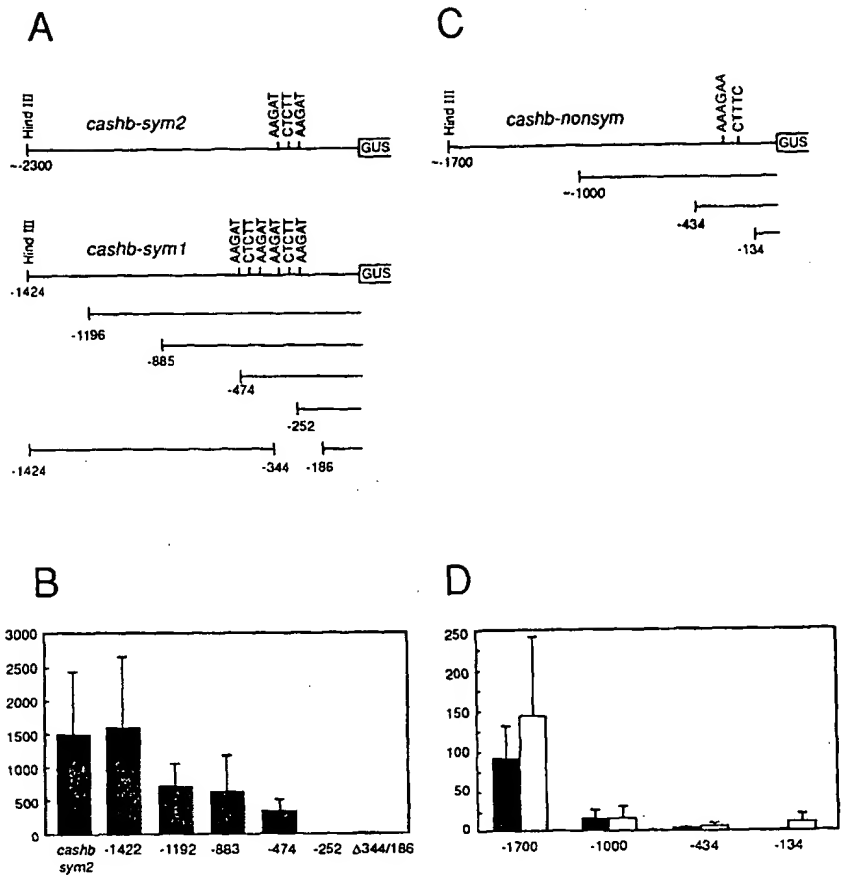

Figure 8. Deletion Analysis of *Casuarina* Hemoglobin Promoters.

(A) Diagrams are shown of the 5′ regions of *cashb-sym1* and *cashb-sym2* promoter–*GUS* constructs together with the 5′ and internal deletions of *cashb-sym1* that were transferred into *Lotus* plants.

(B) GUS activity of *cashb-sym1* in nodules measured by fluorometric assay is given as nanomoles of 4-methylumbelliferone per milligram of protein per hour (vertical axis). Promoter constructs are identified on the horizontal axis.

(C) The 1700-bp *cashb-nonsym* promoter–*GUS* construct (5′ region) together with the 5′ deletions that were introduced into *Lotus* are shown.

(D) GUS activity in nodules (filled columns) and roots (open columns) were measured by fluorometric assay and are given as nanomoles of 4-methylumbelliferone per milligram of protein per hour (vertical axis). Deletion end points are given as nucleotides from the translation initiator ATG. The 95% confidence intervals are indicated above the columns.

as indicated. The digests were electrophoresed on a 0.8% agarose gel. Blotting and hybridization were according to Sambrook et al. (1989). The symbiotic hemoglobin probe was the Hb-SymA cDNA Insert, and the nonsymbiotic hemoglobin probe was a 900-bp HindIII fragment extending from the third exon downstream in the *cashb-nonsym* gene (Christensen et al., 1991).

RNA Gel Blot Analysis

Total RNA was isolated from *Casuarina* nodules, leaves/stems, and roots of uninfected plants by a modification of the method of Hughes and Galau (1988). Equal amounts of RNA, as determined by $OD_{260\ nm}$,

Table 1. Amino Acid Homology between Plant Hemoglobins

Plant Hemoglobin	*Casuarina* hb-nonsym[a] (%)	*Casuarina* hb-sym (%)	*Parasponia* hb (%)
Soybean lba[b]	43	44	40
Parasponia hb[c]	80	52	
Casuarina hb-sym[d]	53		

[a] Christensen et al. (1991).
[b] Hyldig-Nielsen et al. (1982).
[c] Ellfolk (1972); Landsmann et al. (1986).
[d] Kortt et al. (1988).

of each sample were subjected to electrophoresis on a 1.5% formaldehyde agarose gel, blotted, and hybridized. ^{32}P-labeled RNA probes were the Hb-SymA cDNA and a 600-bp genomic HpaI-HindIII fragment containing the third exon of the *cashb-nonsym* gene (Christensen et al., 1991).

Isolation of Genomic Clones

A genomic library in λEMBL4 was made from 15- to 20-kb size-fractionated DNA of a partial Sau3A digest of *Casuarina* DNA (Sambrook et al., 1989). Approximately 240,000 plaque-forming units were screened with a ^{32}P-labeled Hb-SymA cDNA probe. Eight genomic clones were isolated and characterized by restriction mapping. A fragment containing the complete coding region of *cashb-sym1* and a 1.45-kb promoter region (EMBL accession number L28826) was subcloned into pUC118 and pUC119 and sequenced using the Applied Biosystems automatic sequencing system and standard sequencing procedures. A 2.3-kb genomic fragment containing the promoter of a second gene, *cashb-sym2* (EMBL accession number X77693), was subcloned, and 450 bp of the proximal promoter was sequenced.

Construction of Chimeric β-Glucuronidase Genes

The 1.45-kb fragment containing the *cashb-sym1* promoter and the 2.3-kb fragment containing the *cashb-sym2* promoter were cloned into the pALTER vector. A BglII restriction site was generated by mutation at the translation initiator ATG using the Promega in vitro mutagenesis kit. The mutated *cashb-sym1* promoter was sequenced to ensure that the mutation was present. The mutated promoters were cloned using the HindIII-BglII sites into HindIII-BamHI of the pIV20 vector (Hansen et al., 1989). The 5′ deletion series was generated using the double-stranded nested deletion kit from Pharmacia, and end points were determined by sequencing before cloning the promoter deletions as BglII end-filled EcoRI fragments into the BamHI end-filled SalI site of the pIV20 vector. The *cashb-nonsym* promoter used is ~1.7 kb long and was cloned as an XbaI-BglII fragment into the XbaI-BamHI site of the vector pIV20. Using the following restriction sites present in the promoter, 5′ promoter deletions were generated: BglII, SpeI, and XhoI sites that were ligated to BamHI, XbaI, and SalI sites of the pIV20 vector, respectively.

Plant Transformation

Gene constructs were transferred into *Agrobacterium rhizogenes* as described by Van Haute et al. (1983). The AR12 strain carrying the cauliflower mosaic virus 35S–chloramphenicol acetyltransferase gene construct in the left T-DNA segment (Hansen et al., 1989) was used as the vector throughout this study. *Lotus corniculatus* was transformed, regenerated, and nodulated as described previously (Stougaard et al., 1986, 1987; Petit et al., 1987).

Biochemical Assays

Activity from the β-glucuronidase gene was measured in seven to 12 plants of each construct by fluorometric assay (Jefferson et al., 1987). Protein levels in extracts were determined by the dye binding assay of Spector (1978). Chloramphenicol acetyltransferase assays for transformation were performed as described previously (Stougaard et al., 1986).

ACKNOWLEDGMENTS

We thank Carol Andersson and Danny J. Llewellyn for helpful advice and stimulating discussions and Cyril A. Appleby for the gift of *Casuarina* hemoglobin antibodies and for helpful argument. The research was supported by the Carlsberg Foundation, the Danish Research Academy of Science, the Danish Biotechnology Programme, and the Human Frontier Science Program.

Received August 31, 1994; accepted December 20, 1994.

REFERENCES

Appleby, C.A. (1974). Leghemoglobin. In The Biology of Nitrogen Fixation, A. Quispel, ed (Amsterdam: North-Holland Press.), pp. 521–554.

Appleby, C.A. (1984). Leghemoglobin and *Rhizobium* respiration. Annu. Rev. Plant Physiol. **35**, 443–478.

Appleby, C.A., Bogusz, D., Dennis, E.S., and Peacock, W.J. (1988). A role for hemoglobin in all plant roots? Plant Cell Environ. **11**, 359–367.

Bogusz, D., Appleby, C.A., Landsmann, J., Dennis, E.S., Trinick, M.J., and Peacock, W.J. (1988). Functioning hemoglobin genes in non-nodulating plants. Nature **331**, 178–180.

Christensen, T., Dennis, E.S., Peacock, W.J., Landsmann, J., and Marcker, K.A. (1991). Hemoglobin genes in non-legumes: Cloning and characterization of a *Casuarina glauca* hemoglobin gene. Plant Mol. Biol. **16**, 339–344.

Davidowitz, E.J., Creissen, G., Vincze, E., Kiss, G.B., and Lang-Unnasch, N. (1991). Sequence analysis of alfalfa (*Medicago sativa*) leghemoglobin cDNA and genomic clones. Plant Mol. Biol. **16**, 161–165.

de Billy, F., Barker, D.G., Gallusci, P., and Truchet, G. (1991). Leghemoglobin gene transcription is triggered in a single cell layer in the indeterminate nitrogen-fixing root nodules of alfalfa. Plant J. **1**, 27–35.

Elfolk, N. (1972). Leghaemoglobin, a plant haemoglobin. Endeavour **31**, 139–142.

Fleming, A.I., Wittenberg, J.B., Wittenberg, B.A., Dudman, W.F., and Appleby, C.A. (1987). The purification, characterization and ligand-binding kinetics of hemoglobin from root nodules of the non-leguminous *Casuarina glauca-Frankia* symbiosis. Biochim. Biophys. Acta **911**, 209–220.

Gallusci, P., Dedieu, A., Journet, E.P., Huguet, T., and Barker, D.G. (1991). Synchronous expression of leghemoglobin genes in *Medicago truncatula* during nitrogen-fixing root nodule development and response to exogenously supplied nitrate. Plant Mol. Biol. **17**, 335–349.

Gibson, Q.H., Wittenberg, J.B., Wittenberg, B.A., Bogusz, D., and Appleby, C.A. (1989). The kinetics of ligand binding to plant hemoglobins: Structural implications. J. Biol. Chem. **264**, 100–107.

Gilles-Gonzales, M.A., Ditta, G.S., and Helinski, D.R. (1991). A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti.* Nature **350,** 170–172.

Go, M. (1981). Correlation of DNA exonic regions with protein structural units in hemoglobin. Nature **291,** 90–92.

Hansen, J., Jørgensen, J.-E., Stougaard, J., and Marcker, K.A. (1989). Hairy roots—a short cut to transgenic nodules. Plant Cell Rep. **8,** 12–15.

Hughes, D.W., and Galau, G. (1988). Preparation of RNA from cotton leaves and pollen. Plant Mol. Biol. Rep. **6,** 253–257.

Hyldig-Nielsen, J.J., Jensen, E.Ø., Paludan, K., Wiborg, O., Garrett, R., Jørgensen, P., and Marcker, K.A. (1982). The primary structures of two leghemoglobin genes from soybean. Nucl. Acids Res. **10,** 689–701.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusion: β-Glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6,** 3901–3907.

Jeffreys, A.J. (1981). Recent studies on gene evolution using recombinant DNA. In Genetic Engineering 2, R. Williamson, ed (New York: Academic Press), pp. 1–48.

Jensen, E.Ø., Paludan, K., Hyldig-Nielsen, J.J., Jørgensen, P., and Marcker, K.A. (1981). The structure of a chromosomal leghemoglobin gene from soybean. Nature **291,** 667–679.

Kortt, A.A., Inglis, A.S., Fleming, A.I., and Appleby, C.A. (1988). Amino acid sequence of hemoglobin I from root nodules of the nonleguminous *Casuarina glauca-Frankia* symbiosis. FEBS Lett. **231,** 341–346.

Landsmann, J., Dennis, E.S., Higgins, T.J.V., Appleby, C.A., Kortt, A.A., and Peacock, W.J. (1986). Common evolutionary origin of legume and non-legume plant hemoglobins. Nature **324,** 166–168.

Lauridsen, P., Franssen, H., Stougaard, J., Bisseling, T., and Marcker, K.A. (1993). Conserved regulation of the soybean early nodulin Enod2 gene promoter in determinate and indeterminate transgenic root nodules. Plant J. **3,** 483–492.

Medberry, S.L., Lockhart, B.E.L., and Olszewski, N.E. (1992). The Commelina yellow mottle virus promoter is a strong promoter in vascular and reproductive tissues. Plant Cell **4,** 185–192.

Metz, B.A., Welters, P., Hoffman, H.J., Jensen, E.O., Schell, J., and de Bruijn, F.J. (1988). Primary structure and promoter analysis of leghemoglobin genes of the stem-nodulated tropical legume *Sesbania rostrata*: Conserved coding sequences, *cis*-elements and *trans*-acting factors. Mol. Gen. Genet. **214,** 181–191.

Miao, G.-H., and Verma, D.P.S. (1993). Soybean nodulin-26 gene encoding a channel protein is expressed only in the infected cells of nodules and is regulated differently in roots of homologous and heterologous plants. Plant Cell **5,** 781–794.

Nap, J.-P. (1988). Nodulins in Root Nodule Development. PhD Dissertation (Wageningen, The Netherlands: Agricultural University).

Nap, J.-P., and Bisseling, T. (1990). Nodulin function and nodulin gene regulation in root nodule development. In Molecular Biology of Symbiotic Nitrogen Fixation, P.M. Gresshoff, ed (Boca Raton, FL: CRC Press Inc.), pp. 181–229.

Petit, A., Stougaard, J., Kuhle, A., Marcker, K.A., and Tempe, J. (1987). Transformation and regeneration of the legume *Lotus corniculatus*: A system for molecular studies of symbiotic nitrogen fixation. Mol. Gen. Genet. **207,** 245–250.

Ramlov, K.B., Laursen, N.B., Stougaard, J., and Marcker, K.A. (1993). Site directed mutagenesis of the organ-specific element in the soybean leghemoglobin lbc3 promoter. Plant J. **4,** 577–580.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Sandal, N.N., Bojsen, K., and Marcker, K.A. (1987). A small family of nodule-specific genes from soybean. Nucl. Acids Res. **15,** 1507–1519.

Schmülling, T., Schell, J., and Spena, A. (1988). Single genes from *Agrobacterium rhizogenes* influence plant development. EMBO J. **7,** 2621–2629.

Schwintzer, C.R., and Tjepkema, J.D. (1990). The Biology of *Frankia* and Actinorhizal Plants. (New York: Academic Press).

Spector, T. (1978). Refinement of the Coomassie blue method of protein quantification. Anal. Biochem. **86,** 142–146.

Stougaard, J., Marcker, K.A., Otten, L., and Schell, J. (1986). Nodule-specific expression of a chimaeric soybean leghemoglobin gene in transgenic *Lotus corniculatus.* Nature **321,** 669–674.

Stougaard, J., Sandal, N.N., Grøn, A., Kuhle, A., and Marcker, K.A. (1987). 5′ analysis of the soybean leghemoglobin lbc3 gene: Regulatory elements required for promoter activity and organ specificity. EMBO J. **6,** 3565–3569.

Stougaard, J., Jørgensen, J.-E., Christensen, T., Kühle, A., and Marcker, K.A. (1990). Interdependence and nodule-specificity of *cis*-acting regulatory elements in the soybean leghemoglobin lbc3 and N23 gene promoters. Mol. Gen. Genet. **220,** 353–360.

Szabados, L., Ratet, P., Grunenberg, B., and de Bruijn, F.J. (1990). Functional analysis of the *Sesbania rostrata* leghemoglobin *glb3* gene 5′-upstream region in transgenic *Lotus corniculatus* and *Nicotiana tabacum* plants. Plant Cell **2,** 973–986.

Szczyglowski, K., Szabados, L., Fujimoto, S.Y., Silver, D., and de Bruijn, F.J. (1994). Site-specific mutagenesis of the nodule-infected cell expression (NICE) element and the AT-rich element (ATRE-BS2*) of the *Sesbania rostrata* leghemoglobin *glb3* promoter. Plant Cell **6,** 317–332.

Taylor, E.R., Nie, X.Z., MacGregor, A.W., and Hill, R.D. (1994). A cereal hemoglobin gene is expressed in seed and root tissues under anaerobic conditions. Plant Mol. Biol. **24,** 853–862.

Van Haute, E., Joos, H., Maes, M., Warren, G., Van Montagu, M., and Schell, J. (1983). Intergenetic transfer and exchange recombination of restriction fragments cloned in pBR322: A novel strategy for reversed genetics of the Ti-plasmids of *Agrobacterium tumefaciens.* EMBO J. **2,** 411–417.

Welters, P., Metz, B., Felix, G., Palme, K., Szscyglowski, K., and de Bruijn, F.J. (1993). Interaction of a rhizobial DNA-binding protein with the promoter region of plant leghemoglobin genes. Plant Physiol. **102,** 1095–1107.

Wittenberg, J.B., Bergersen, F.J., Appleby, C.A., and Turner, G.L. (1974). Facilitated oxygen diffusion: The role of leghemoglobin in nitrogen fixation by bacteroids isolated from soybean root nodules. J. Biol. Chem. **249,** 4057–4066.